The Landmark Approach Applied to Implied Volatility

av

Andres Sørensen Susrud

MASTEROPPGAVE

for graden

Master in Modelling and Dataanalysis

(Master of Science)

Det matematisk- naturvitenskapelige fakultet
Universitetet i Oslo

November 2012

1 Preface

This master thesis is a part of my masters degree in Modelling & Data Analysis, Finance, Insurance and risk at University of Oslo.

During this writing process I have spent some time in Munich, Germany, and in this context I want to thank Dr. Thilo Meyer-Brandis and Dr. Alex Langnau from LMU (Ludwig Maximillian Universität) / Allianz Investment management for their support and encouragement.

I would like in this preface to thank the people at LMU /AIM for the fantastic stay in Munich, where I have learned a lot related to practical approach to working towards financial mathematics. Frank Proske at UiO for supporting me in the process of writing the thesis, and interesting topic. A big thank to my girlfriend Helén Schjønhaug and my family for supporting me over the time of my studies.

Contents

List of Figures

List of Tables

List of Algorithms

Part I
Introduction

There exists numerous books on the topic of *technical analysis,* but very few that has a scientific approach to verifying that it is actually anything to it. The *technical analysis* varies from reoccurring patterns, such as double bottom, double top, head-shoulder, and other formations in the price movements and barrier levels known as support and resistance levels. The support and resistance levels may be calculated with many different approaches, and has in the past been drawn manually by looking at charts. This is also the typical presentation in several books.

We will in this thesis look at the intraday DAX future price and the eurodollar spot price, to determine whether there exist any stastistically significant evidence of support and resistance levels. Our approach will not be any of the particular methods for predicting / drawing the lines of support and resistance lines, but instead looking at whether there are levels that are more reoccuring in the DAX future price than in a mimicking Wiener process. If there are any levels with sufficient predictability, one would have an edge in the market and be able to make profits that deviates from the efficient market hyphotesis.

The Thesis is organized in the following way: In Part II we are presenting the smoothing tool that we will be using throughout the thesis. This model is capturing "landmarks" in the timeseries, that would be closely to what a human would look at the charts, and also capturing the global and local minimas and maximas.

In Part III we introduce tools we will use to compare the results in the observed DAX future price or the eurodollar spot price versus the results in a Wiener process. The Shannon Entropy, that will tell us about the clustering of the process. And the statistical tests used in the comparison. Part IV contains the mathematical properties we derived from the landmark model with respect to a standard Wiener process, and chapter V contains the proposed new method for estimating volatility. Part VI is the research of evidence of support and resistance levels and, and Part VII is the concluding remarks.

Part II
Smoothing

To be able to spot the technical indicators easier, a lot of smoothing techniques has been used to remove noise in the charts. Moving average is used in many cases by technical analysis practitioner for determining technical signals[1]. Andrew W. Lo et al. are using the kernel regression smoother[2] in their paper testing for statistical evidence of geometrical formations in U.S. stocks. Due to the lag effect of the Moving Average, I will instead use the Landmark model in this thesis. The landmark model as introduced by Chang-Shing Perng, Haixun Wang, Sylvia R. Zhang, D. Stott Parker[3] will be able to capture every local and global extreme value of the chart, depending on the given chosen resolution.

In technical analysis, extremevalues are the most important, because it is what the technical signals / formations are based on. The traditional Technical Analysis is based on the human interpretation from studying the charts, and drawing lines from between extremevalues, or discovering reoccuring geometrical patterns in the market. From the automated point of view it is easier to deal with the local extreme values from the landmark model, because the technical signals / formations are based on strict algorithmic rules.

2 Landmark model

"The Landmark Model does not follow traditional similarity models that rely on pointwise Euclidean distance. Instead, it leads to Landmark Similarity, a general model of similarity that is consistent with human intuition and episodic memory"[4].

A person is asked to look at a graph for a short period of time, and then reproduce it by memory. If the person memorizes the major turning points, and then connect them, it will make a "good" reproduction of the graph. Varying the resolution in the landmark model will yield a different reproduction of the timeseries. For very large resolution p one would only capture the starting and ending point, and for very small resolution p, one would capture every point in the time series. The inputs in the paper by Perng et. al. are p & d, where p is the minimal change on the y axxis (process), while d is the minimum change in the x axxis (time). The resolution is the treshold value for when a point in the time series is considered to be a landmark (Figure 1). In the paper "Landmarks: A New Model for Similarity-Based Pattern Querying in Time Series Databases", the authors takes both the resolution p with respect to the movement in price, and the resolution of d with respect to time. We will in this thesis only consider the Landmark model with respect to the resolution of p, the change in price as

[1] Reference 2
[2] Reference 7
[3] Reference 1.
[4] Ref[1]

shown in Figure 2. We will also consider absolute value changes with respect to p and not relative changes with respect to p as proposed in the paper by Pern, Wang, Zhang and Parker. This is because we will mainly be using the DAX futures as the testing material, and the minimum movement in the DAX future price is 0.5 points.

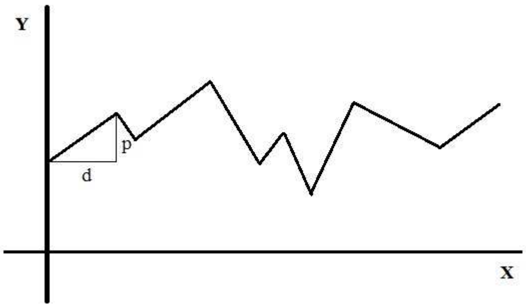


Figure 1: Landmark example
Example on how the landmarks are chosen, given the size of p and d.

Algorithm 1 Altorithm for updating the landmarks

Inputs: $n = 1$, S_t, p

1. for i in 1:T
2. if $n < 2$,$LM_n = S_n$
3. if $LM_n < LM_{n-1}$ & $S_{n+1} < LM_n$ or $LM_n > LM_{n-1}$ & $S_{n+1} > LM_n$. Then $LM_n = S_{n+1}$
4. if $\frac{1}{2}\left|LM_n - LM_{n-1}\right| < p$. Then LM_n is deleted and goes to next iteration.
5. if condition 2,3 and 4 does <u>not</u> hold, a new landmark is added. $LM_{n+1} = S_{n+1}$
6. $n = n + 1$
7. returns LM_n

LM_n is the landmark representation of the timeseries S_t. Example shown in Figure 2 below.

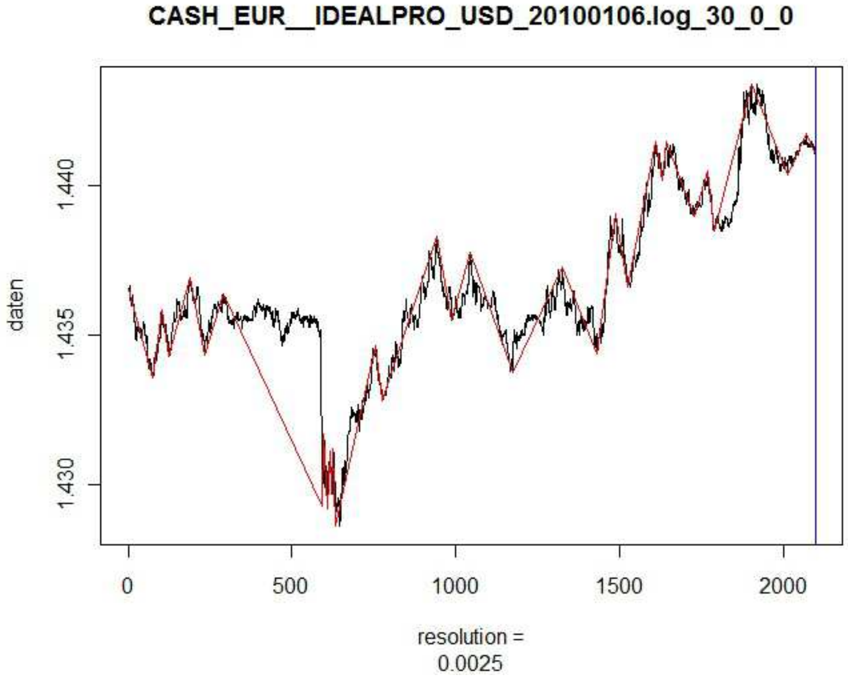


Figure 2: Landmark example FX data

In our case, we have set $d \rightarrow \infty$, so that we won't take any change in time to be considered. This is because, we only want landmarks that is dependent on price, and not on time.

Part III

Tools (Observed data versus Brownian motion)

In order to verify the existence or measure the extent of the technical indicator, we will in this thesis compare it with a Wiener process that posesses the same properties as our observed dataset. There are also other ways of doing this, eg. creating a trading strategy that yields a profit that exceeds the expected return according to the "efficient market hyphothesis". We will later in the thesis also compare the backtest results of a trading strategy based on the technical indicators, but primarily compare the test results to the theoretical values and distributions of the Wiener process.

3 Diffusion model

The term structure we use in this paper is the Wiener process of with standardized volatility $\sigma = 1$. The discussion whether this is the correct choice of model for the market will not be discussed in this paper.

$$DAX = \int_0^T \sigma dB_t$$

$$dDAX = \sigma\, dB_T$$

Where $B_t \sim N(0,t)$ Standard Brownian motion. We will consider in the sequel the diffusion process to have length of the day to be m, and number of steps to be m, so that the $(B_t - B_{t+1})\,\tilde{}N(0,1)$.

4 Brownian motion

4.1 Properties of a Brownian motion[5]:

For a probability space $(\Omega,)\,\mathcal{F}, P)$ let the process be a Brownian motion . Then:

- $B_0(\omega) = 0$
- for all ω, t $B_t(\omega)$ is continuous, or has a continuous version
- for all $0 \leq t_1 \leq t_2 \leq \ldots$, then $B_{t_1}, B_{t_2} - B_{t_1}, B_{t_3} - B_{t_2}, \ldots$ are independent
- $E[B_t(\omega)] = 0$, $\forall t$
- $E\left[(B_t(\omega) - B_s(\omega))^2\right] = t - s$, $\forall\, t,\ s$
- for all t, s $B_t - B_s \sim N(0, (t-s))$
- $\hat{B}_t = \frac{1}{c} B_{c^2 t}$, for $c > 0$ being a constant, is also a Brownian motion (scaling property)

5 Poisson process (for an extension of the model)[6]

The Poisson process is often used as a model for financial time series instead of the Brownian motion. Some of the main reason for this is due to the lack of jumps in the Brownian motion, the lack of fat tails, and also it often captures the empirical observations in a more accurate fashion than the Brownian motion.

[5] Reference 5
[6] Reference 8

5.1 The Poisson process

The poisson process is defined by:

$$N_t = \sum_{n \geq 1} 1_{t \geq T_n} \tag{1}$$

where $(\tau_i)_{i \geq 1}$ is a sequence of independent exponentially distributed random variables with parameter $\lambda > 0$ and $T_n = \sum_{i=1}^{n} \tau_i$.

5.2 Compensated Poisson processes

$$\tilde{N}_t = N_t - \lambda t \tag{2}$$

The new process $\tilde{N}_t$ now follows a centered version of the Poisson law, with characteristic function

$$\Phi_{\tilde{N}_t}(z) = exp\left[\lambda t \left(e^{iz} - 1 - iz\right)\right] \tag{3}$$

N_t does not follow the martingale property, but the new process $\tilde{N}_t$ does,

$$E\left[\tilde{N}_t | \mathcal{F}_s\right] = \tilde{N}_s, \, \forall t > s \tag{4}$$

Where $\mathcal{F}_s$ is the fintration generated from $N_s, 0 \leq s \leq t$. The compensated Poisson process is no longer integervalued, and no longer a counting process. The rescaled version $\frac{\tilde{N}_t}{\lambda}$ has the two first moments equivalent to the standard Wiener process.

$$E\left[\frac{\tilde{N}_t}{\lambda}\right] = 0 \tag{5}$$

$$Var\left[\frac{\tilde{N}_t}{\lambda}\right] = t \tag{6}$$

$\left(\frac{\tilde{N}_t}{\lambda}\right)_{t \in [0,T]} \Rightarrow^{\lambda \to \infty} (W_t)_{t \in [0,T]}$, that is the rescaled version is convergent in distribution to the standardized Wiener process.

5.3 Stable law & properties

A property of Brownian motion is the self similarity property, as in section 3.1: If W is a Wiener process, then

$$\forall a > 0, \left(\frac{W_{at}}{\sqrt{a}}\right)_{t \geq 0} =^d (W_t)_{t \geq 0} \tag{7}$$

A Levy process X_t is said to be selfsimilar if

$$\forall a > 0, \exists b(a) > 0 : \left(\frac{X_{at}}{b(a)}\right)_{t \geq 0} =^d (X_t)_{t \geq 0} \tag{8}$$

Since the characteristic function of X_t has the form $\Phi_{X_t}(z) = \exp[-t\psi(z)]$[7], this property is equivalent to the property of the characteristic function:

$$\forall a > 0,\, \exists b(a) > 0 : \, \Phi_{X_t}(z)^a = \Phi_{X_t}(zb(a)),\, \forall z \tag{9}$$

It can be shown[8] that for every stable distribution there exists a constant $\alpha \in (0, 2]$ such that in equation (15), $b(a) = a^{\frac{1}{\alpha}}$. The constant $b(a)$ is called the index of stability, and these distributions which have this property are also called $\alpha - stable$ distributions. Setting $\alpha = 2$ we see that we will get the Wiener process as in equation (7).

6 Normalization

In order to compare the observed data to a standard Wiener process we must normalize the observed data. The way we have done this in this thesis is simply to define a new process

$$\tilde{S}_t = \frac{S_t - S_{t-1}}{\sigma_{obs}} + \tilde{S}_{t-s} \tag{10}$$

where $\{S_t\}_{t\geq 0}$ is the data point of the Dax future, and σ_{obs} is the observed standard deviation. Now the new process $\tilde{S}_t$ will have standard deviation of 1, and we can compare it with a standard Wiener process.

7 Shannon entropy

The Shannon entropy or *the information entropy* is a measure of uncertainty represented by the discrete probability distribution. The theory was introduced by Claude E. Shannon in his 1948 paper: A Mathematical Theory of Communication[9]. The use of the entropy in this thesis is to test if there is a difference in the entropy of observed market data versus the entropy of a Wiener process with same properties. We expect the entropy in the brownian motion to be higher than in the observed market data case, since the Brownian motion has no memory.

7.1 Definition of Shannon's formula:

$$H = -\sum_{i=1}^{N} p_i \, log_2\,(p_i)\,, \tag{11}$$

where N is the number of states $\# A$, and p_i is the state probability $p_i = pr\,(x \in a_i)$.

[7] Reference 8, chapter 3.7
[8] Reference 8, chapter 3.7
[9] Reference 3

Here
$\sum_{i=1}^{N} p_i = 1$
and

$$\max_H = -\sum_{i=1}^{N} p\, log_2\left(p\right) = -\sum_{i=1}^{N} \frac{1}{N} log_2\left(\frac{1}{N}\right) = log_2\left(N\right) \tag{12}$$

where $p_i = p\,\forall i \in [1, N]$ and $p = \frac{1}{N}$

$$\min_H = -\sum_{i=1}^{N} p_i\, log_2\left(\frac{1}{p_i}\right) = 0 \tag{13}$$

where $p_i = \begin{cases} 0 & , i = j \\ 1 & , i \neq j \end{cases}$

This tells us that the lower the entropy of the distribution is, the more we know of the clustering or probability of the different subsets $a_i \subseteq A$. In the $\max_H$ all the states as equally likely. We are in this paper normalizing the entropy by dividing the calculated entropy by the $\max_S$ and will therefore operate with $S \in [0, 1]$. This is to be able to compare the entropy of observations of different N.

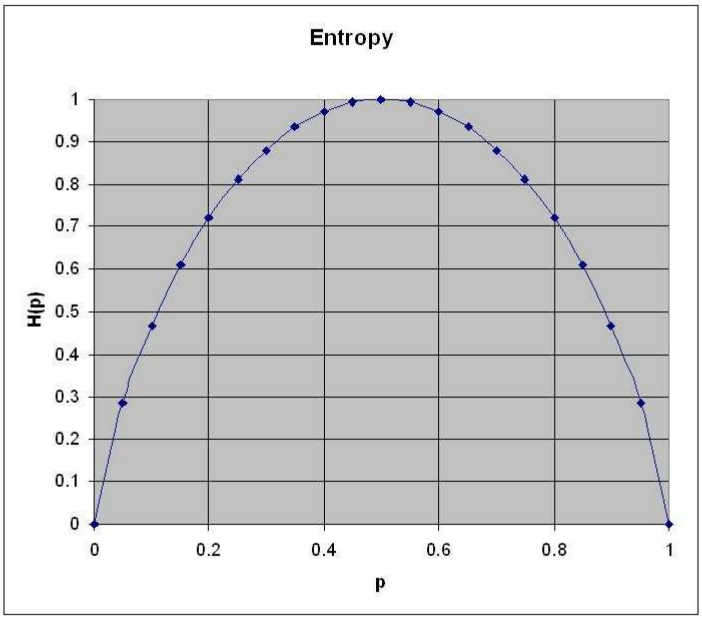


Figure 3: Entropy distribution
Graphical explanation of the H with respect to the probability distribution of a fair coin toss. In this for 2 states, and we see that $p_i = \frac{1}{2}$,$i \in [1, 2]$ gives the highest Entropy. For example a fair coin would have entropy 1, while an unfair coin will have lower entropy.

8 Statistics

I will in the following sections compare the results of the studies, mainly the difference between the observed results and the benchmark results. I will mostly be using descriptive statistics to show the results. I will in this section briefly introduce the techniques applied.

8.1 Basics

8.1.1 Mean

The mean or the sample mean is the most useful measure for the centre. The measure is the arithmetic average over the samples. For independent identically distributed (i.i.d) random variables the sample mean is also the *expected value*, due to the central limit theorem. This property is often used for Monte Carlo simulation.

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{N} x_i$$

8.1.2 Standard deviation

The standard deviation is a measure of the variability of the samples. The classical way of estimating this is looking at the squared distance between the sample and the mean of the sample. The standard deviation is also what in finance is referred to as the volatility.

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

Unbiased estimator for standard deviation:

$$\hat{\sigma}_X = \sqrt{\frac{1}{n-1}\sum_{i=1}^{N}(x_i - \bar{x})^2}$$

Standard deviation of the sample mean:

$$\sigma_{\bar{X}} = \frac{\sigma_x}{\sqrt{N}}$$

8.2 Confidence Intervals

The confidence intervall will tell us not only the central measure (mean), but also the uncertainty of the measure. The confidence intervall will give us an invervall where the central measure (mean) will be located. The extent of the uncertainty is due to the α level. Often is α chosen such that we are 95 % certain that the mean is within the desired area. eg

$$p\left(-\mathcal{Z}_{\frac{\alpha}{2}} < \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < \mathcal{Z}_{\frac{\alpha}{2}}\right) = 100\,(1-\alpha)\,\%,$$

where $\bar{X}$ is the desired area. and $\mathcal{Z}_{\frac{\alpha}{2}}$ is the quantile of the standard normal distribution with respect to the α[10].

$$p\left(-\mathcal{Z}_{\frac{\alpha}{2}} * \sigma/\sqrt{n} < \bar{X} - \mu < \mathcal{Z}_{\frac{\alpha}{2}} * \sigma/\sqrt{n}\right)$$

$$p\left(\bar{X} - \mathcal{Z}_{\frac{\alpha}{2}} * \sigma/\sqrt{n} < \mu < \bar{X} + \mathcal{Z}_{\frac{\alpha}{2}} * \sigma/\sqrt{n}\right)$$

which gives us the formula for calculating the confidence intervall for the mean.

$$\bar{X} \pm \mathcal{Z}_{\frac{\alpha}{2}}\sigma_x\sqrt{n} \tag{14}$$

This formula will be used as a confidence intervall for mean values in the Monte Carlo simmulations later.

[10] Reference 4 chap. 9.4

8.3 Two sample t-test

Given we want to test the difference,$X-Y$ of two different samples X and Y. If X and Y are of different length, and variance, we must use the pooled variance estimate for the two samples.

Assumption:

1. X and Y are independent
2. X are normally distributed
3. For large samples 2 will be normal distributed by C.L.T

$$T = \frac{\bar{X} - \bar{Y} - (\mu_x - \mu_y)}{\sqrt{\frac{S_x^2}{n_x} + \frac{S_y^2}{n_y}}}$$

where μ_x is the mean of the X samples, μ_y mean of the Y samples as in 7.1. n_x is the sample size of X, and n_y is the sample size of Y. S_x^2 is the unbiased standard deviation estimator squared from section 7.1.2, and S_y^2 is the unbiased standard deviation estimator squared.

$$\nu = \frac{\left(\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}\right)}{\frac{(s_x^2/n_x)^2}{n_x-1} + \frac{\left(s_y^2/n_y\right)^2}{n_y-1}}$$

Then T is approximately t-distributed by ν degrees of freedom.

8.4 Hypothesis testing

Testing for eg. differences in two variables, two populations. We reject the proposed hypothesis, or null hypothesis, often denoted H_0, by 2 different styles in the t-test. There is a one sided t test, where we check whether the H_0 : $\mu_1 - \mu_2 < \delta$ or $\mu_1 - \mu_2 > \delta$. In the two sided test: $H_0 : \mu_1 - \mu_2 \neq \delta$. We reject if the test observator t:

- $t \geq t_{\alpha,\nu}$, for $\mu_1 - \mu_2 > \delta$
- $t \leq t_{\alpha,\nu}$, for $\mu_1 - \mu_2 < \delta$
- $t \geq t_{\frac{\alpha}{2},\nu}$ or $t \geq t_{\frac{\alpha}{2},\nu}$, for $\mu_1 - \mu_2 \neq \delta$

8.5 Central limit theorem[11]

Let $X_1, X_2, \ldots, X_n$ be a random vector with mean μ and standard deviation σ. Then in limit as $n \to \infty$, the standardized version of $\bar{X}$ and T_0 have the standard normal distribution:

$$\lim_{n\to\infty} P\left(\frac{\bar{X} - \mu}{\sigma\sqrt{n}} \leq z\right) = P\left(Z \leq z\right) = \Phi\left(z\right) \tag{15}$$

[11] Reference 4

and

$$\lim_{n \to \infty} P\left(\frac{T_0 - n\mu}{\sqrt{n}\sigma} \leq z\right) = P\left(Z \leq z\right) = \Phi\left(z\right), \tag{16}$$

where Z is a standard normal random variable, and $T_0 = X_1 + X_2 + \ldots + X_n$ with expectation $E[T_0] = n\mu$ and $\sigma_{T_0} = \sqrt{n}\sigma$[12]

Part IV
Estimating historical volatility from the observed number of landmarks

9 Estimating historical volatility from the observed number of landmarks

The idea behind this method is to be able to estimate daily volatility, by comparing the number of landmarks for different resolutions in the landmark model (sec. 3) for a Wiener process with known standard volatility, with the number of landmarks in the observed dataset with the resolution p. By changing the resolution p as input in the landmark algorithm, we will obtain less landmarks with a bigger resolution p, and more landmarks with a smaller resolution p. This is also intiutive as the “treshold” as the minimum move required to be accepted is the size of p. We have also calculated analytically the number of landmarks, given the resolution p, and a constant that gives the number of landmarks of resolution $p = 1$. (In the bordercase of this estimator it will take only open and close / first and last datapoint of the time series)

9.1 Analytical properties of $N_p(t)$

We define the $N_p(t)$ to be the number of landmarks with resolution p and time $t = [0, t]$, and $X_t = N_1(t)$ to be the number of landmarks given resolution $p = 1$. We then define 2 rules according to the properties of a Brownian motion:

$$N_{\sqrt{\lambda}p}\left(\lambda t\right) = N_p\left(t\right) \tag{17}$$

$$N_p\left(\lambda t\right) = \lambda N_p\left(t\right) \tag{18}$$

$$N_p(t) = \lambda N_{\sqrt{\lambda}p}(t) \tag{19}$$

Equation (17) is the Brownian scaling law, and equation (18) is the Markov property, such that the expected number of landmark in a process that is λ times longer, will have λ times more landmarks. Equation (19) is (17) and (18)

[12] Reference 4, chap 6.2

combined. By the Brownian scaling law we also have that $\lambda N_{\sqrt{\lambda}p}(t) = \lambda^2 N_{\lambda p}(t)$. this gives us (20)

$$N_p(t) = \lambda^2 N_{\lambda p}(t) \tag{20}$$

We define now:

$$X_t = \lambda N_{\sqrt{\lambda}}(t) = \lambda^2 N_\lambda(t) \tag{21}$$

$$N_\lambda(t) = \frac{X_t}{\lambda^2}, \tag{22}$$

and we define a new variable λ' such that

$$N_{\lambda'}(t)\,\lambda'^2 = N_\lambda(t)\,\lambda^2 \tag{23}$$

$$N_{\lambda'}(t) = \left(\frac{\lambda}{\lambda'}\right)^2 N_\lambda(t) \tag{24}$$

If we now substitute λ' with p and λ with 1, we get:

$$N_p(t) = \left(\frac{1}{p}\right)^2 X_t \tag{25}$$

We now take the log of equation (25) and get:

$$\ln N_p = -2\ln p + \ln X_t \tag{26}$$

Hence the slope of the $\ln N_p = -2\ln p$ as we also see from the numerical example:

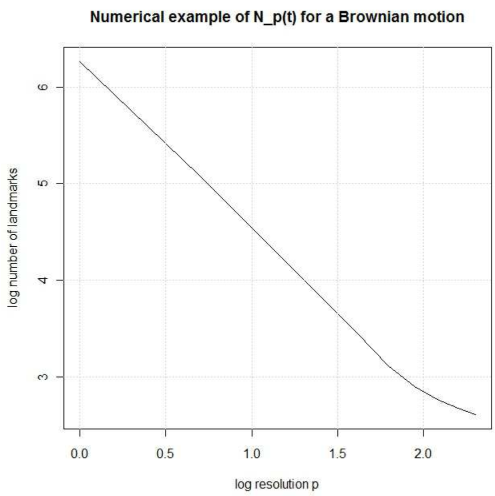


Figure 4: Numerical slope of $N_p(t)$

We see however that the slope is declining for large p. This is due to the length of the Wiener process, and that for large p the dataset is not long enough to get the theoretical number of landmarks. The same also occurs for small $p's$ and this is because the $\Delta t = (t_{i-1} - t_i)$ is to big, and that the path has several more landmarks between t_{i-1} and t_i.

9.2 Making the benchmark

We are now simulating a standardized Wiener process as described in section 2, with a benchmark volatility, and m number of time steps from $[0, t]$.

$$sd\left(B_{t+1} - B_t\right) = 1$$

For different resolutions $p' = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11\}$, we count the number of landmarks for the process. $N_{p'}(t)$ =number of landmarks with resolution p', within the time interval $[0, t]$. We introduce λ as a fixed constant. $N_{p,\sigma}(t)$ is the number of landmarks given the resolution p for a Wiener process with unknown volatility σ. $N_{p,1}(t)$ is the number of landmarks given resolution p for a Wiener process with standardized volatility of $\sigma = 1$.

9.3 Calculating the implied resolution $p^{implied}$

From the conditions

$$\begin{aligned} N_{p,\sigma}(t) &= N_{p,1}\left(t\sigma^2\right) && (27) \\ &= N_{\frac{p}{\sigma},1}(t) && (28) \end{aligned}$$

$N_{p,\sigma}(t)$ is the number of landmarks for a process with volatility σ and resolution p. We can now establish the estimation equation:

$$N_{p',1}^{Benchmark}(t) = N_{p,\sigma}^{Observed}(t) \tag{29}$$

We now define a new variable $p^{implied}$ that is the resolution on the observed number of landmarks with unknown volatility σ, such that $N_{p^{implied},1}^{Benchmark}(t) = N_{p,\sigma}^{Observed}(t)$. From equation (18), (19) and (20) we get:

$$p^{implied} = \frac{p^{Benchmark}}{\sigma} \tag{30}$$

Using now the $N_{p',1}(t)$ for all the $p' = \{1, \ldots, 11\}$,and matching the $p^{implied}$ such that $N_p^{Benchmark}(t) = N_p^{Observed}(t)$, we will get an estimate of σ for each $p' > p^{implied}$. The final estimate of the volatility is now the mean of the estimates we get from the implied resolution $p^{implied}$.

$$\hat{\sigma} = \frac{1}{11}\sum_{i=1}^{11}\sigma_i \tag{31}$$

Algorithm 2 Algorithm for calculating the implied resolution $p^{implied}$ for the process Z_t

inputs: $p^{Benchmark}, N_{p^{Benchmark},1}(m), m =$ (number of steps in benchmark process), Z_t, l (number of steps in process Z_t)

1. $n = \sum_{i=1}^{N} \mathcal{X}_{\{p_i^{obs} \leq p_1^{BM}\}}$
2. for i in $p^{Benchmark}$(n,n+1,...,11)
3. calculate $N_{p^{Benchmark},\sigma}(l)$ for Z_t
4. Scale the observed $N_{p^{Benchmark},\sigma}(l)$ by $\frac{m}{l}$
5. $\log p_i^{implied} = \log(p_i^{Benchmark}) + \frac{(\log N_i^{Obs} - \log N_i^{Benchmark})}{(-2)}$
6. $\hat{\sigma}_i = \exp \frac{\log p_i^{Benchmark}}{\log p_i^{implied}}$
7. $\hat{\sigma} = \sum_{i=n}^{11} \hat{\sigma}_i$
8. return $\hat{\sigma}$

In the example below is the corresponding resolution to p in the observed dataset. p' can be observed in the graph below and the corresponding table. The $p^{implied}$ is as in the table below calculated by linear interpolation between the observed points $p = \{1, \ldots, 11\}$, to match the observed N_{p_i}.

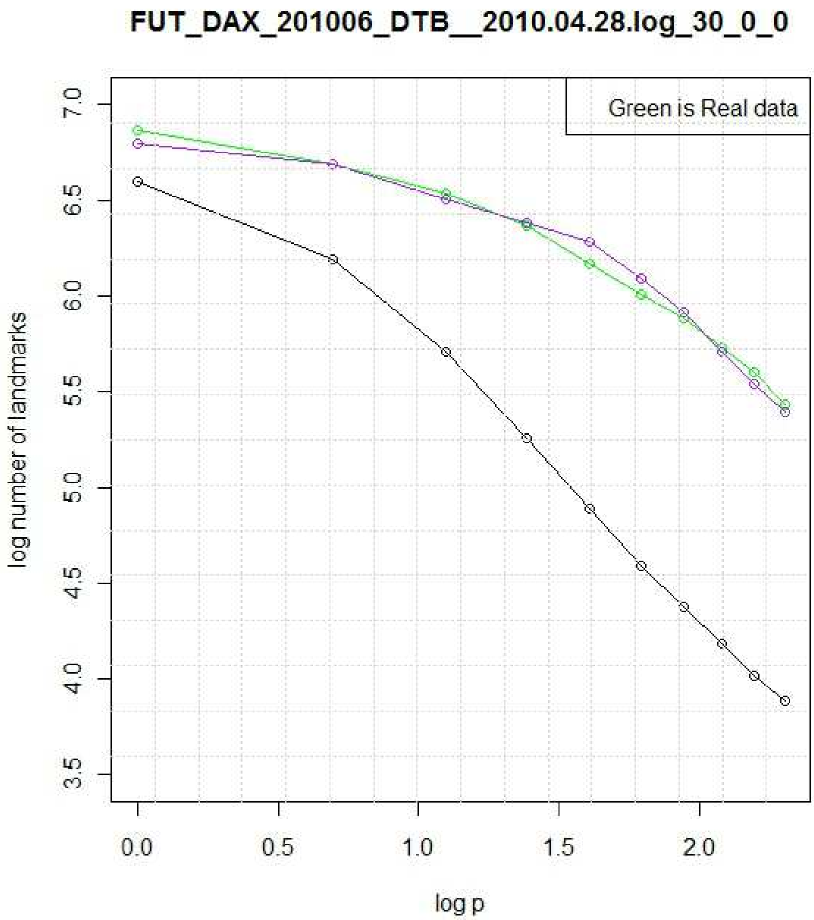


Figure 5: Example implying volatility
The black is the Log number of landmarks in the Brownian motion simulation with standard volatility. Green is observed data of the DAX, and purple is a new Brownian motion generated with the estimated volatility $\hat{\sigma}$.

log p_i	p'	$\frac{exp(p_i)}{exp(p')}$
1.099	0.097	2.722
1.386	0.396	2.691
1.609	0.710	2.459
1.792	0.849	2.567
1.946	0.953	2.698
2.079	1.083	2.709
2.197	1.168	2.798

Table 1: Example implying volatility

9.4 Correcting for "non-observed" landmarks

We have in Figure 4. seen that the theoretical slope of the log number of landmarks should be as in equation (17): $\ln N_p = -2 \ln p + \ln X_t$. Because the Brownian motion is continuous, nowhere differential and self similar, we must correct the observed numbers of landmarks by a scaling factor due to the "missing" landmarks in between Δt.

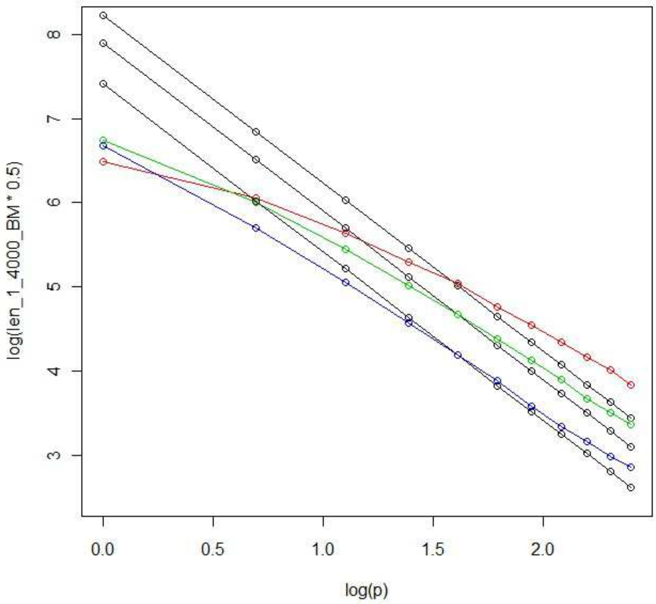


Figure 6: Theoretical Benchmark vs. Montecarlo

We see that the slope of the log number of landmarks in the Monte Carlo simulated Wiener process varies for different Δt, and also specificly varies from the theoretical slope as calculated in equation (17).

The benchmark log number of landmarks is calculated by taking the mean of the number of landmarks for a Monte Carlo simulated Wiener process with standardized volatility, for resolution $p = 5$. We derive the curve from this with the theoretical slope from equation (17). We therefore know that for $p = 5$, the Monte Carlo simulated Wiener process will have the same number of landmarks as in the benchmark. We establish an equation that will correct the deviation with respect to the known good fit ($p = 5$), and we call this p_0, for the known Δt_0 and the known σ_0.

We derive the lambda from the tripple Δt_0, p_0, σ_0, where we know the Monte Carlo has a good fit with the theoretical benchmark $N_p(t)$.

$$\lambda = \frac{1}{\Delta t_0}\left(\frac{p_0}{\sigma_0}\right)^2 \tag{32}$$

Hence an appropriate choice for Δt_p given p, σ_o, λ is

$$\Delta t_p = \frac{1}{\lambda}\left(\frac{p}{\sigma_0}\right)^2 \tag{33}$$

We are undercounting the number of landmarks for for $p < p0$, since we are missing the landmarks in between Δt_0 for smaller p. We are now adjusting the Δt with respect to p, given σ_0 and the constant λ. Now $N_{p,\Delta t_0}(t)$ is the number of landmarks with Δt_0. We want to represent the number of landmarks for a fixed Δt_0, given the adjustment in equation (3) for $p \neq p_0$. We do this by adjusting the

$$N_{p,\sqrt{\Delta t_0}}(t/\Delta t_0)\, for\, p \neq p_0 \tag{34}$$

with the adjustment $\Delta t_p(3)$. we have from the brownian scaling law:

$$W_t = c * W_{t*\frac{1}{c^2}} \tag{35}$$

applying this to $N_{p,\Delta t_p}$ give us:

$$N_{p,\sqrt{\Delta t_p}*\sqrt{\frac{\Delta t_0}{\Delta t_p}}}\left(\frac{t}{\Delta t_p}\left(\frac{1}{\frac{\Delta t_0}{\Delta t_p}}\right)\right) = N_{p,\Delta t_0}\left(t\sqrt{\frac{\Delta t_0}{\Delta t_p}}\right) = \sqrt{\frac{\Delta t_0}{\Delta t_p}} N_{p,\Delta t_0}\left(t/\Delta t_0\right) \tag{36}$$

And we have the scaling factor

$$F = \sqrt{\frac{\Delta t_0}{\Delta t_p}} \tag{37}$$

if we insert equation (2) and (3) we get.

$$N_{p,\frac{1}{\lambda}\left(\frac{p}{\sigma_0}\right)^2}(t) = N_{p,\frac{1}{\lambda}\left(\frac{p}{\sigma_0}\right)^2 * \frac{\frac{1}{\lambda}\left(\frac{p_0}{\sigma_0}\right)^2}{\frac{1}{\lambda}\left(\frac{p}{\sigma_0}\right)^2}}\left(t * \sqrt{\frac{\frac{1}{\lambda}\left(\frac{p_0}{\sigma_0}\right)^2}{\frac{1}{\lambda}\left(\frac{p}{\sigma_0}\right)^2}}\right) \tag{38}$$

Now the final formula for $N_{p,\Delta t}\,(t)$is:

$$N_{p,\Delta t} = F * N_{p,\Delta t_0},\, \forall p \tag{39}$$

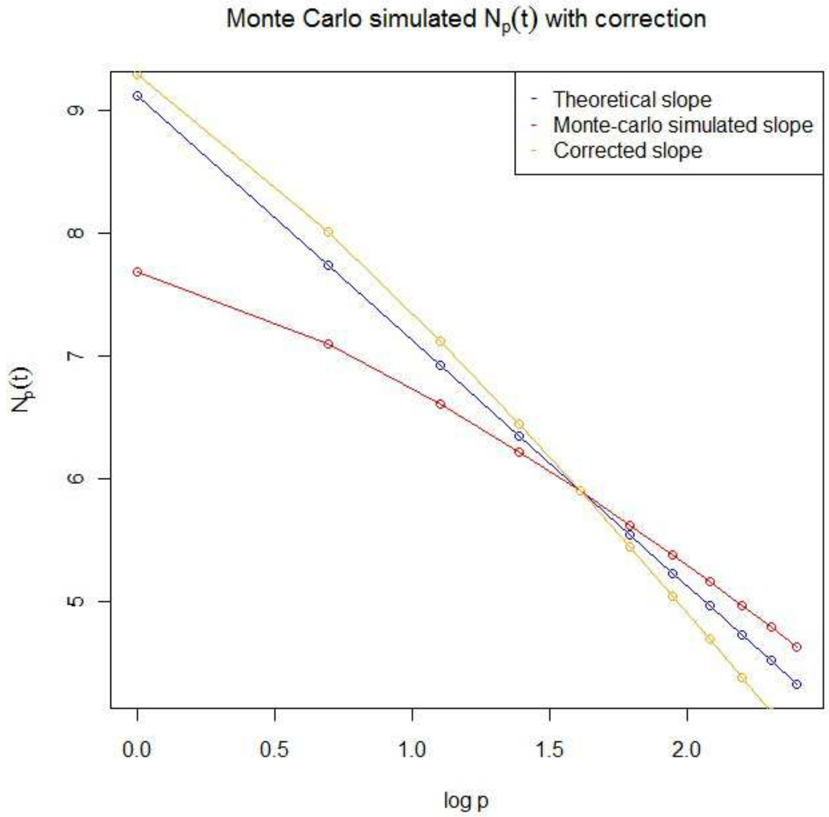


Figure 7: Monte Carlo simulated $N_p\,(t)$ before and after correction
Change to the one with more paths & error band

9.5 Extensions

A possible generalization to the landmark volatility estimator is considering the more general $\alpha - stable$ Levy process from section 4.3, since the self-similarity property is the one of the main ideas behind the volatility estimator. Another possible extension is to use the landmarks as a levy measure estimator.

9.6 Jump size / intensity estimator

In order to test the jump size / intensity estimator, we must compare it to a Wiener process. The Wiener process has no discontinuities and therefore will also the time scaled difference between the landmarks be smaller than if we have discontinuities.

9.7 Normalizing landmarks

In an attempt to detect jumps in a timeseries, we will normalize the landmarks of different resolution p with respec to the time difference between the landmarks. Using this method we will get higher normalized value for local minimas or maximas that are due to change in a short period of time. We define now the normalized landmarks as

$$L\tilde{M}_n = \frac{LM_{n+1}^{t_{n+1}} - LM_n^{t_n}}{(t_{n+1}) - t_n} \tag{40}$$

9.8 Detecting outerliers in the $L\tilde{M}_n$

There are several ways of detecting outerliers among these is Grubb's test, Dixon's test & Cochron's C test. We will test the different outliers tests, and then match it against a generated Levy Jump diffusion. We will be using the Grubb's test[13] via the R package 'outliers' to test if the biggest differences in the time normalized Landmarks are outliers.

9.9 Jump frequency / frequency of outliers

From the outlier test, we find the number of outliers that are eg. 95% significant. We then assume that the jump frequency is normally distributed and collect the jump frequency for the in example last 20 days, and get a frequency distribution from this.

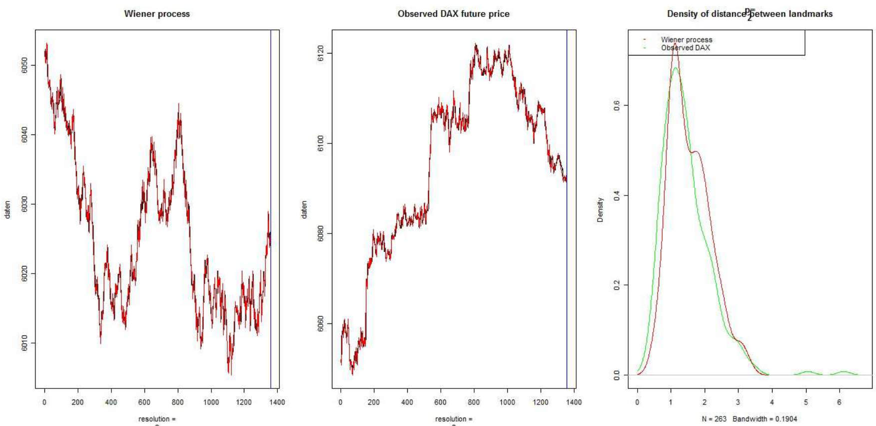


Figure 8: Density of distance between normalized landmarks

The example shows the Wiener process of standard deviation 1, an observed normalized DAX future price, and the corresponding density of the distances between the landmarks. It seems some of the distances are outliers, and could be counted as jumps.

[13] Reference 9

In determining if the outliers are significant, we use the Grubbs test and test with $\alpha = 0.05$.

```
> grubbs.test(diff.lm)

    Grubbs test for one outlier

data:  diff.lm
G = 6.4043, U = 0.8429, p-value = 3.283e-09
alternative hypothesis: highest value 6.12313889965518 is an outlier

> which.max(diff.lm)
[1] 32
> grubbs.test(diff.lm[-32])

    Grubbs test for one outlier

data:  diff.lm[-32]
G = 5.4059, U = 0.8876, p-value = 3.485e-06
alternative hypothesis: highest value 5.06201600081476 is an outlier

> which.max(diff.lm[-32])
[1] 107
> grubbs.test(diff.lm[-c(32,108)])

    Grubbs test for one outlier

data:  diff.lm[-c(32, 108)]
G = 3.1958, U = 0.9606, p-value = 0.1645
alternative hypothesis: highest value 3.44681599470101 is an outlier
```

Figure 9: Grubbs test of landmarks distance

We detect two outliers in the difference in the landmarks for the observed process. The third largest distance is not an outlier due to it's p-value of 0.1645, and we're rejecting at 0.05. This is what we intuitively saw in the density plot (figure 8)

9.9.1 Algorithm for testing the number of outliers

We want to make this an automated procedure, and therefore we will recursively count the number of outlier, by eliminating the biggest, and then run the test again, as shown in figure 8.

Algorithm 3 Recursive outlier detection

1. inputs Landmarks, outlier, m=1
2. Calculate the distances between the landmarks, and normalize with respect to time between landmarks
3. Run Grubb's test for outliers
4. If outlier is detected by $\alpha = 0.05$, $outlier_m = max(diff^{normalized}_{landmark})$, $m = m + 1$, remove and go to step 3
5. If Grubb's test for outliers returns a p-value greater than $\alpha = 0.05$, return $outlier$

The number of outliers returned is the now the estimated number of jumps for this observed dataset. We can make an estimator for the jump frequency by running algorithm 3 on a sequence of days, and then take the expected value of the number of outliers.

$$\lambda = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{41}$$

where x_i is the number of outliers in day i, and n is the number of days in a sequence.

10 Efficiency of the estimator

10.1 classical estimation

$$\hat{\sigma}_{DAX} = \frac{1}{n-1}\sum_{i=1}^{n}\left(log\left(\frac{S_{i+1}}{S_i}\right)\right)^2 \tag{42}$$

$$\hat{\sigma}_{DAX} = \sqrt{\frac{1}{n-1}\sum_{i=i}^{n}\left((S_{t+1}-S_t)-(\overline{S_{t+1}-S_t})\right)^2} \tag{43}$$

where $(\overline{S_{t+1}-S_t}) = \frac{1}{n}\sum_{i=1}^{n}(S_{t+1}-S_t)$. The classical estimator is biased in the sense that $E\left[\hat{\sigma}_{DAX} =\right] = \sigma$, where $E\left[\,\right]$ denotes taking the expectation. And so the variance of the estimator $VAR\left(\hat{\sigma}_{DAX}\right) = \sigma^4\left(\frac{2}{n-1}+\frac{\kappa}{n}\right)$, where n is the sample size and κ is the kurtosis.

10.2 Efficiency

A way to determine the efficiency is proposed by Garman & Klass [14] is to define

$$eff\left(\hat{y}\right) = \frac{var\left(\hat{\sigma}_0^2\right)}{var\left(\hat{y}\right)} \tag{44}$$

where $\hat{\sigma}_0^2$ is the classical way of estimating volatility, and $\hat{y}$ is the proposed way. Higher $eff\left(\hat{y}\right)$ indicates a smaller variance in the proposed volatility estimator.

[14] ref. 2

10.2.1 Comparison of proposed volatility estimator to classical approach

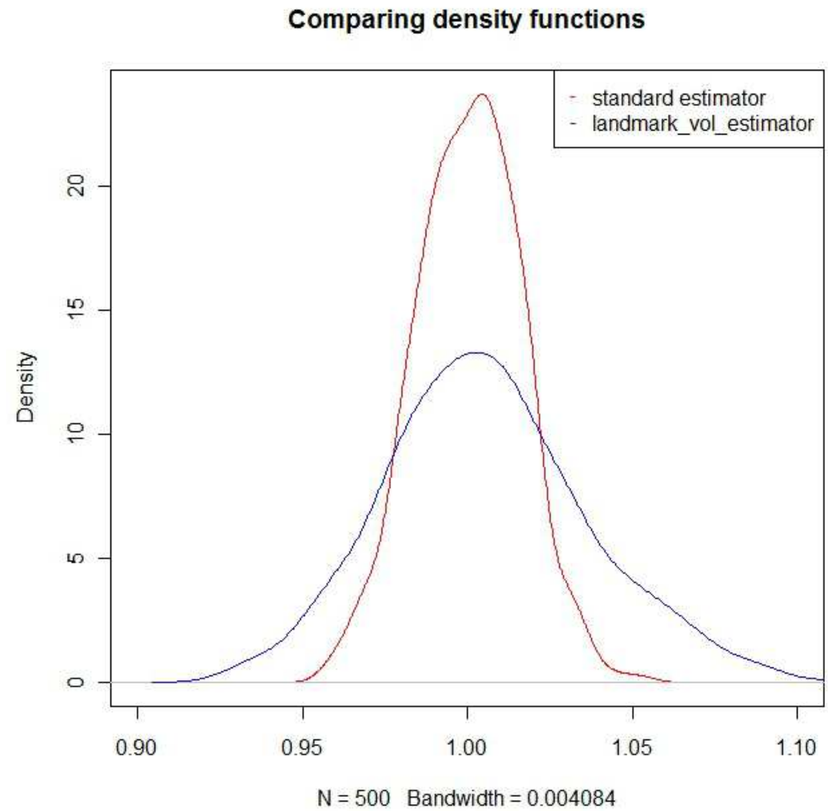


Figure 10: Comparison new estimator - classical estimator

10.3 Conclusion

From the efficiency measure in equation (44)

Part V
Technical indicators

11 Support and resistance levels

The construction of support and resistance levels seems to be more of a psychological interpretation of the market than scientific. The support and resistance levels are considered to be barriers that the security process has problems passing, due to psychological eg. round numbers, previous highs, lows or pattern effects such as double bottom, double top etc. Support levels are levels below the current price that the market price is more likely to bounce back from rather than fall through. Resistance levels is the opposite, the resistance level is above the price process and if reached, expected to bounce back down, rather than rise above. As the price process breaks through the support or resistance level, the level changes from eg. being support to being resistance. Eg. a price process that went below the support level of 1.36 is believed to have problems breaking above this level again, so the level 1.36 is now considered a resistance level.

The most common calculation method predicting support and resistance levels are pivot point calculation, Fibonacci retracements, Bollinger bands. Different kind of moving average techniques are also considered support and resistance lines in some cases.[15]

12 Band probability

We first want to look at whether there are any parts of the daily movement of the observed price process that is more frequently visited. We're doing this by splitting up the normalized daily movements in B parts from daily low to daily high. We then calculate the band probability of each band by:

$$\{p_i\}_{i \leq B} = P\left(LM_i \in B_i\right) \tag{45}$$

for different resolution p in the landmark algorithm. If there are certain bands that have a higher probability in the observed time series compared to the Wiener process with the same properties, it may be an indicator that there exist a support or resistance level within the band. By increasing the number of bands, we are tuning in on where potentially the support or resistance lines would be, and we would expect a higher band probability for the bands where the support or resistance lines would be within. Black line is probability distribution for the observed DAX futures time series, and red is the probability distributions for mimicking Wiener processes.

[15] Reference 2.

Algorithm 4 Calculating the discrete probability distribution

1. Create the landmarks using the Algorithm 1.
2. Split up the sample space of the time series from low to high, in B number of bands.
3. Count number of landmarks within each section, and divide by the number of total landmarks for the time series.

In the figures below we have averaged the probability bands over our total dataset of normalized observed DAX futures, to see if there are any bands, that have a systematic difference from the Wiener process of the same properties.

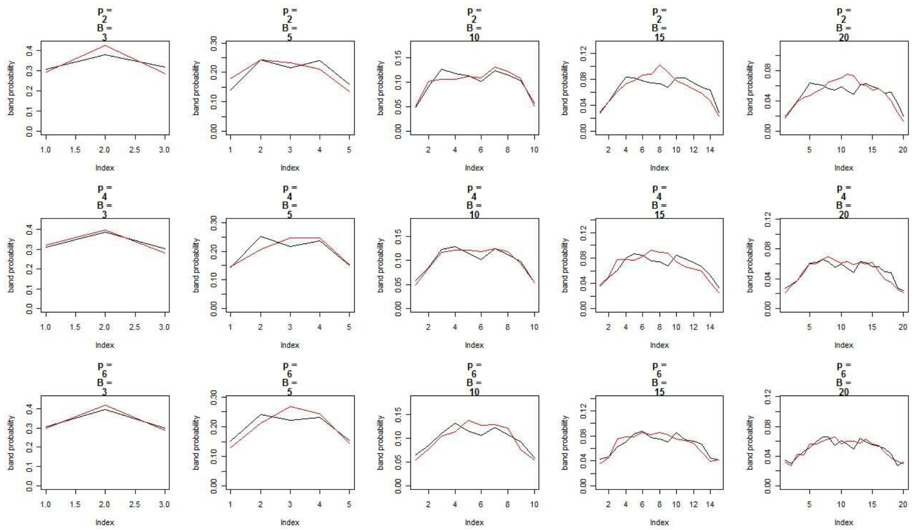


Figure 11: Band probability I
Shows the band probability for resolution 2, 4 and 6, according the the number of bands B. Band 1 indicates the lowest range of the price, and the highest B indicates the highest range of the price. Black line is the band probability distribution for the observed timeseries, and red line the band probability distribution for the Wiener process.

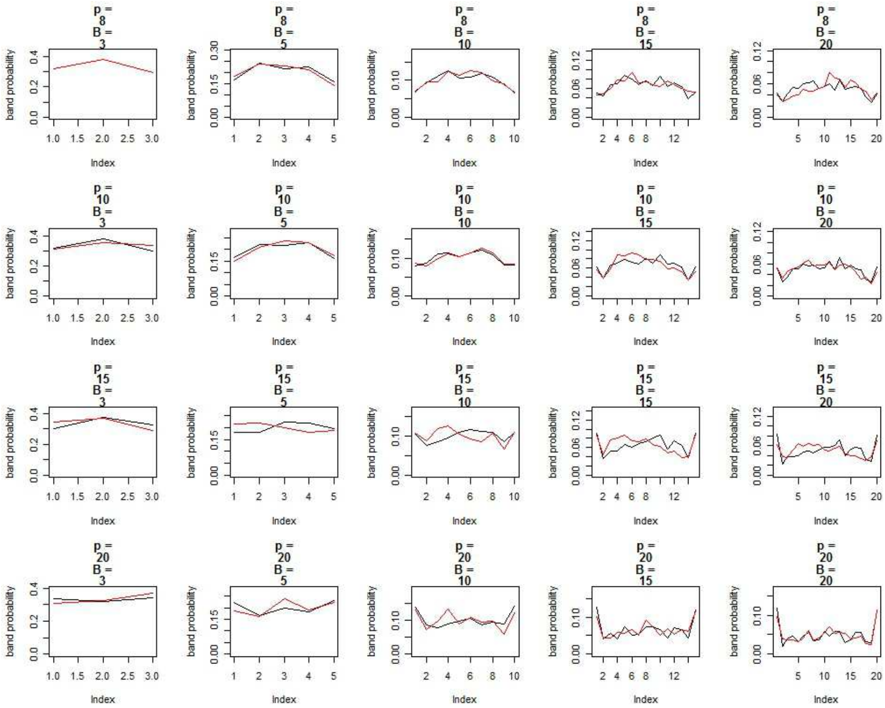


Figure 12: Band probability II
Shows the band probability for resolution 8, 10, 15 and 20, according the the number of bands B. Band 1 indicates the lowest range of the price, and the highest B indicates the highest range of the price. Black line is the band probability distribution for the observed timeseries, and red line the band probability distribution for the Wiener process.

By looking at the plots above, we get no clear impression whether there exist support or resistance levels. It seems though that for smaller p, (Figure 5) that the highest bands seems to have a systematic higher probability for the observed process than the Wiener process.

13 Entropy test

In the previous section we calculated the discrete probability distribution for the landmarks in the observed processes and in the Wiener processes. We want to look for, if there are any clustering of the landmarks by calculating the entropy of the discrete probability distribution. The entropy, as introduced in section 4. will tell us how "fair" the discrete probability distributions are. The lower the entropy, the more clustered distribution of landmarks are within this band. If the entropy of the observed discrete probability distributions are significantly lower than in the Wiener process, we may have an indication that in some bands, there exist a support or resistance level. This tells us that there are levels that has reoccuring prices significantly deviates from random behaviour

We can calculate the entropy by two different approaches. First by calculating the entropy of the averaged discrete probability distributions, as plotted in previous section. Or we can calculate the entropy for each observed normalized daily DAX future time series, and then average the entropy values over the number of days. In this section we will be doing the latter by calculating the probabilities according to algorithm 4.

$$\{p_i\}_{1 \leq B} = pr\left(LM_i \in B_i\right)$$

And then calculating the entropy H with formula (11). Results are given in tables below.

Observed								
B / P	1	2	4	6	8	10	15	20
3	0.8874	0.8972	0.9045	0.9086	0.9084	0.9132	0.9108	0.9026
5	0.8784	0.8876	0.8990	0.9063	0.9097	0.9141	0.9110	0.8775
10	0.8803	0.8904	0.9026	0.9087	0.9074	0.9020	0.8657	0.7897
15	0.885	0.8947	0.9046	0.9056	0.8946	0.8790	0.8177	0.7311
20	0.8899	0.8990	0.9061	0.9000	0.8831	0.8590	0.7837	0.6895

(a) Entropy Observed

Wiener process								
B / P	1	2	4	6	8	10	15	20
3	0.9149	0.9197	0.9201	0.9187	0.9231	0.9240	0.9239	0.9235
5	0.9107	0.9125	0.9139	0.9142	0.9189	0.9223	0.9196	0.9023
10	0.9137	0.9140	0.9186	0.9220	0.9219	0.9156	0.8875	0.8306
15	0.9167	0.9185	0.9186	0.9233	0.9141	0.9039	0.8509	0.7760
20	0.9197	0.9215	0.9235	0.9187	0.9054	0.8871	0.8168	0.7358

(b) Entropy Wiener process

Table 2: Entropy tables

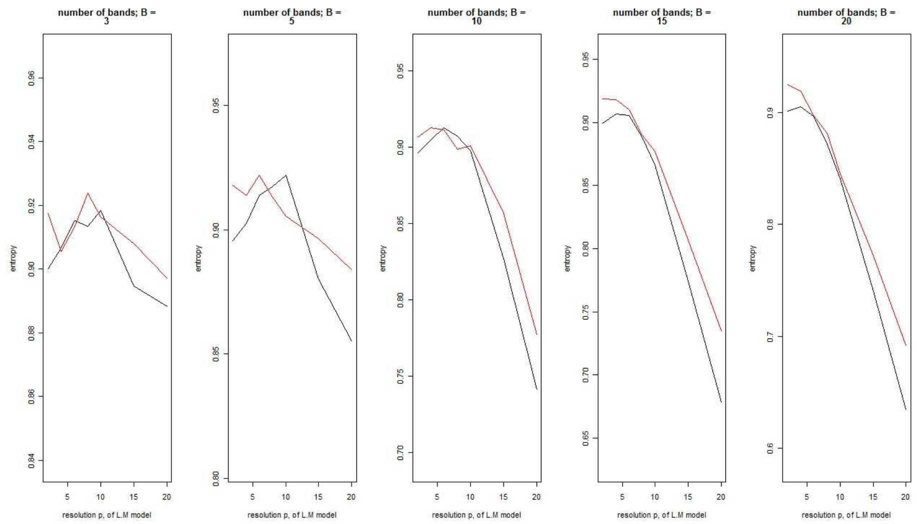


Figure 13: Entropy Observed vs. Wiener process

The red lines in the plot is the entropy for the Brownian motion, while the black line is the entropy for the observed market data. The reason for the big becay in the end seems to be that the higher the resolution of the P, the less landmarks we will get to represent the dataset. Now these will also just be in a few bands, and therefore many of the bands will have zero probability, and hence a lower entropy.

13.1 two-sample t-test for entropy

We want to test the difference in the entropy of the observed data, versus the entropy of the Wiener process. This is done by a two-sample t-test as described in section 6.3. We would now reject the null hypothesis that there are no difference between the entropies for the observed dataset μ_2 and the entropy for the Wiener process μ_1, $\mu_1 - \mu_2 = 0$, if $t \geq t_{\frac{\alpha}{2},\nu}\, or\, t \geq t_{\frac{\alpha}{2},\nu},\, for\, \mu_1 - \mu_2 \neq 0$. The degrees of freedom needed for the test statistic is higher than 120, so the $t_{\alpha/2,\nu}$ is equal to the standard normal $\mathcal{Z}_{\alpha/2}$, since we have simmulated the paths of the wiener process with $n = 1000$, and the number of observed days are 427.

t-statistic for $\mu_1 - \mu_2$, $t_{\alpha/2,\nu} = \{1.96, 2.576\}$for $\alpha = \{0.025, 0.005\}$								
B / P	1	2	4	6	8	10	15	20
3	3.928	2.903	2.474	1.602	2.307	1.704	1.841	2.574
5	5.793	4.565	2.958	1.574	1.805	1.619	1.475	3.244
10	7.399	5.357	3.728	3.096	3.228	2.916	3.765	5.447
15	7.562	5.811	3.300	4.263	4.425	5.419	5.917	6.388
20	7.817	6.038	4.656	4.779	5.391	6.429	6.201	6.871

Table 3: two-sample t-test table for the entropy

We see that a few values in the table has a t-statistic that is lower than the $t_{\alpha/2,\nu}$, and we can therefore not reject the hypothesis that there is no difference between observed and Wiener process. What we also see is that for $B \geq 10$, we have strong significant evidence that there is a difference between μ_1 and μ_2. This is supporting the theory that there may be certain levels in the market where the price is more likely to have turning points within (support & resistance levels)

14 Testing for support and resistance levels in observed data

We want in this thesis to test if there exist certain levels in the timeseries that are significantly reoccurring. By running the landmark algorithm for different resolutions, we will get a number of landmarks that represents the original time series. We have in the previous section shown that the entropy of the normalized observed **DAX futures price (EURUSD futures price)** is statistical significantly lower than for a Wiener process with the same properties. This does not however tell us where the specific clustering is about, or whether it is possible to predict the levels where the clustering would be about. We will in this section test different "areas" of the sample space of the time series, to see if there are any levels that are significantly deviating from the mimicking Wiener process. If there are strong deviations from the Wiener process, it may be an indication that the efficient market hypothesis is violated, and we can exploit this deviation in a trading strategy.

14.1 Daily minimum and maximum value as support and resistance

By testing the number of reoccurences in the daily minimum and maximum values with an error margin, we are testing for whether the intraday previous high and intraday previous low will present itself as possible support and resistance levels. We are running a Monte Carlo simulation for a standardized Wiener process and counting the number of reoccurences to the landmark with the maximum and minimum values, and it's corresponding 95% confidence interval, for different resolution p, and different error margin. Then we are calculating the landmarks in the normalized **DAX futures price time series** and counting the number of reoccurences to the maximum and minimum values.

14.1.1 EURUSD

The figures below is the number of reoccurences to daily minimum and daily maximum for the euro-dollar spot price. The Monte Carlo simulation for the Wiener process is not done with volatility = 1 but the average volatility of the observed dataset. The length of the Wiener process is also the average of the length of the observed euro-dollar prices process.

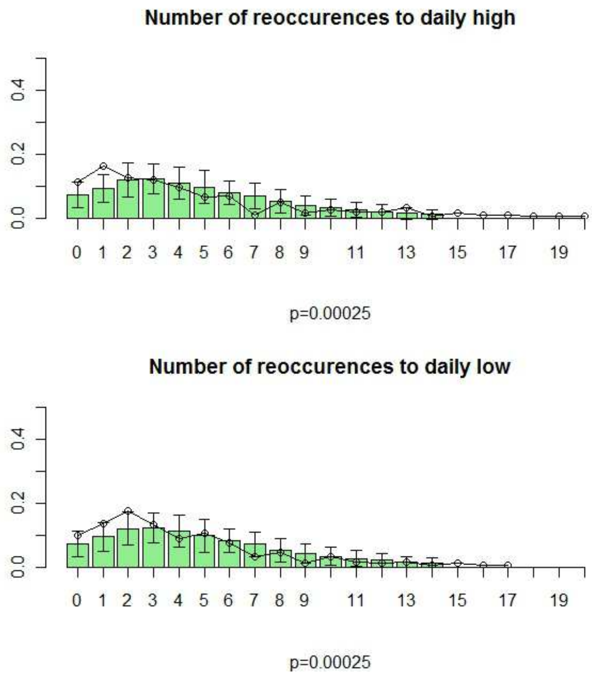


Figure 14: Number of reoccurences MIN / MAX p = 0.00025

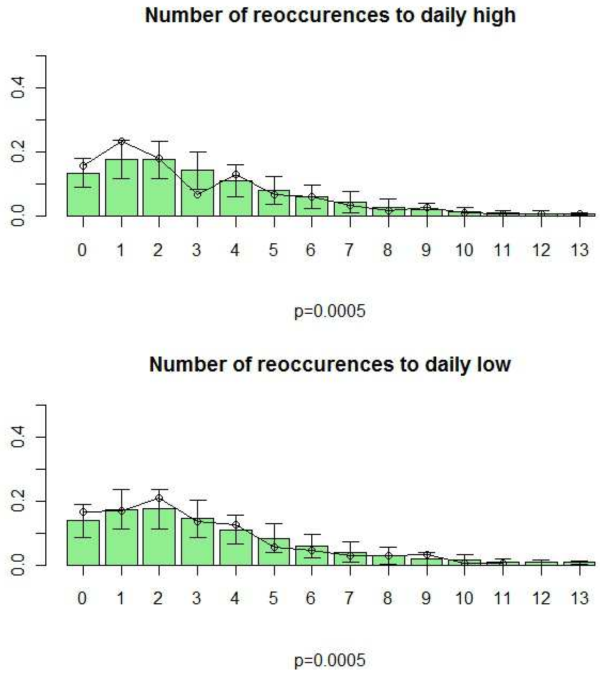


Figure 15: Number of reoccurences MIN / MAX p = 0.0005

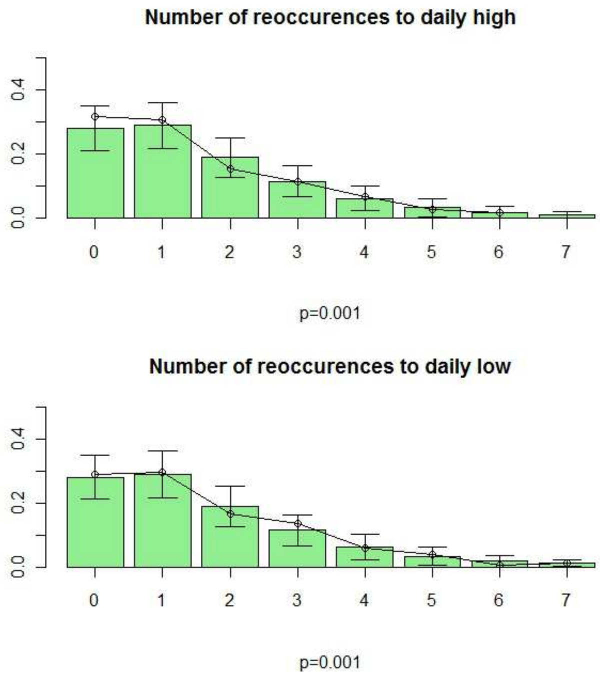
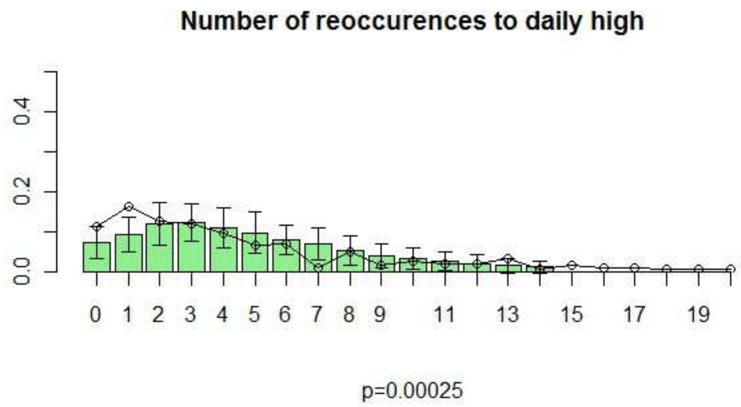


Figure 16: Number of reoccurences MIN / MAX p = 0.001

14.1.2 DAX Future

The figures below is the number of reoccurences to the daily high and daily low for the DAX futures price. The Monte Carlo simulation for the Wiener process is done with the average length of the observed data set and with standardized vol = 1.



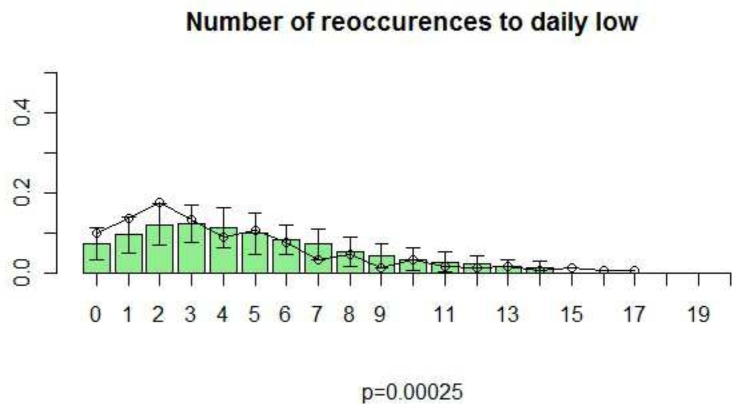


Figure 17: Number of reoccurences MIN / MAX p = 1
The y-axis shows the probability for x number of reoccurences.

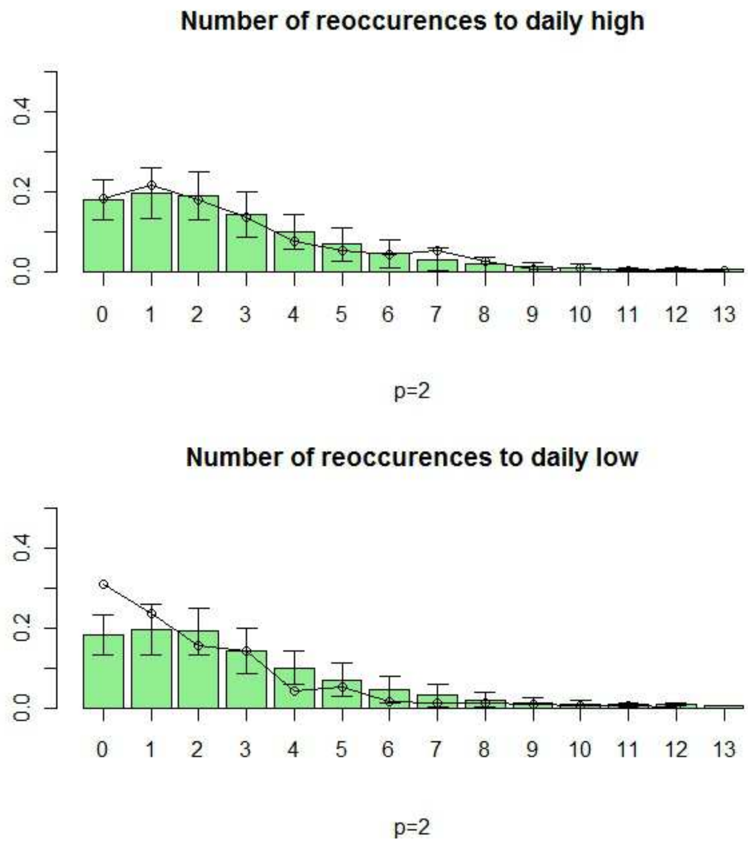
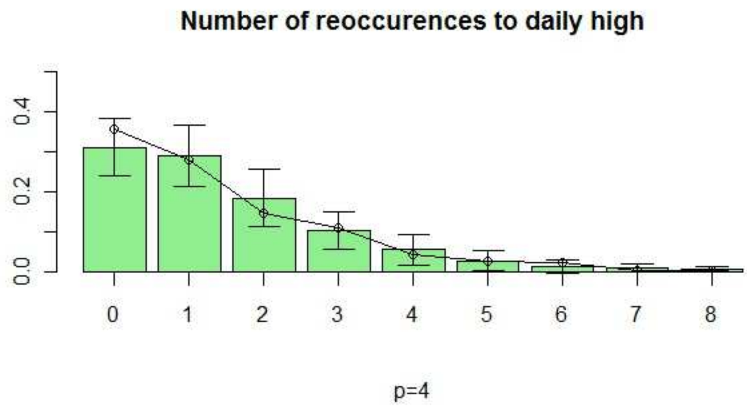


Figure 18: Number of reoccurences MIN / MAX p = 2
The y-axis shows the probability for x number of reoccurences.



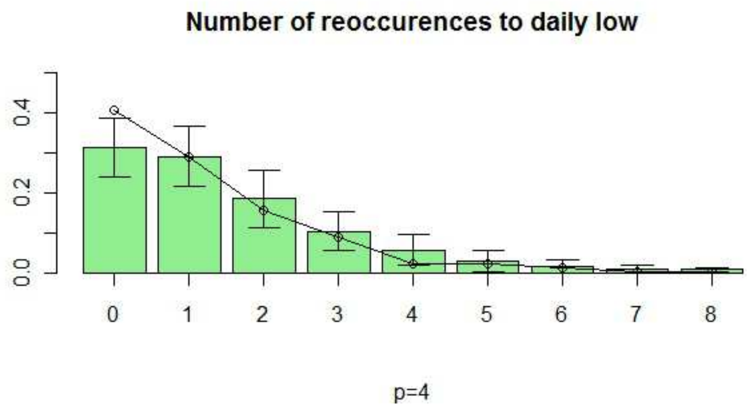


Figure 19: Number of reoccurences MIN / MAX p = 4
The y-axis shows the probability for x number of reoccurences.

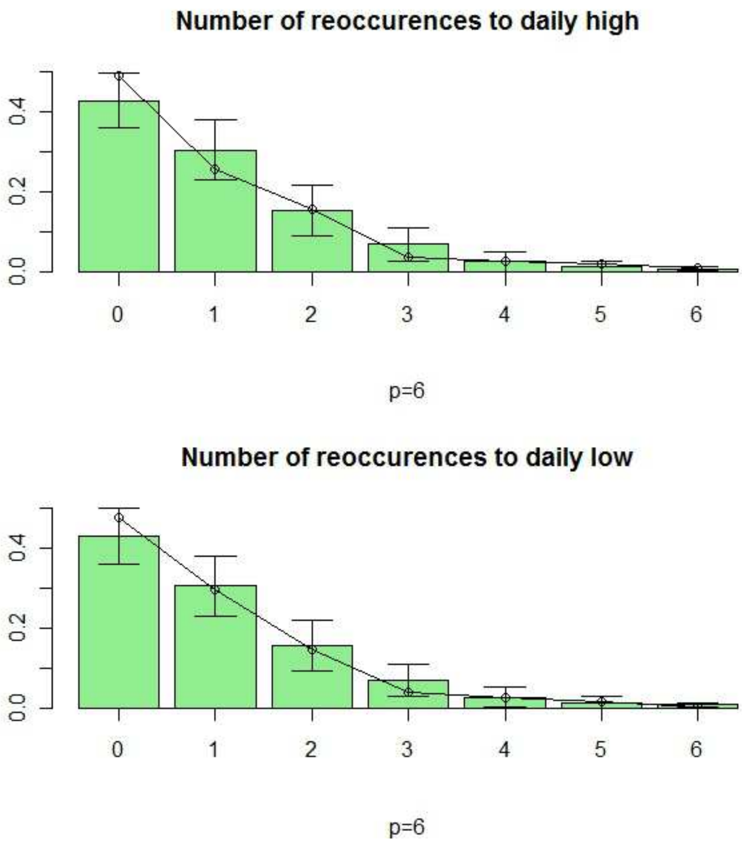


Figure 20: Number of reoccurences MIN / MAX p = 6
The y-axis shows the probability for x number of reoccurences.

14.2 Sloping support and resistance lines

Sloping support and resistance lines, are levels that has a growth coefficient rather than a flat level, as examined in the previous subsection. The sloping support and resistance lines are as vertical support and resistance lines expected to be hit more often in the observed data than in the Wiener process we are comparing it against.

Algorithm 5 Calculating sloping support and resistance lines

1. Calculate the landmarks for the time series
2. inputs: LM_p,LM_t, bw (errorband width)
3. for n in 2: $\#LM * \frac{1}{4}$
4. $slope = \frac{LM_p - LM_{p-1}}{LM_t - LM_{t-1}}$
5. $num_{LM} = \#LM_p \in LM_p + (slope * t) \pm bw$
6. return $max(num_{LM}, slope)$

The comparison in this section is done by taking the average length / number of increments in the observed paths we have. We are then normalizing the observed data set with respect to it's volatility. And generating a standard Wiener process of the same length. To distinguish between support level and resistance level, we have testet for landmarks that are local minimas, and local maximas. We then consider the local minimas to hit the support level and bounce back from the sloping support level. Likewise for local maximas, we consider them to touch the sloping resistance level, and bounce back. A discrete probability is then calculated by dividing the number of landmarks within the slope, on the total number of landmarks for the observed time series, and similar for the Wiener process.

mean probability of reoccurence to sloping resistance level								
p (resolution)	1	2	4	6	8	10	15	20
Observed	0.182	0.179	0.178	0.18	0.183	0.181	0.159	0.109
Wiener process	0.166	0.164	0.168	0.171	0.171	0.170	0.158	0.115

(a) DAX futures

mean probability of reoccurence to sloping support level								
p (resolution)	1	2	4	6	8	10	15	20
Observed	0.174	0.170	0.168	0.169	0.166	0.162	0.127	0.076
Wiener process	0.159	0.157	0.160	0.159	0.161	0.155	0.160	0.101

(b) Wiener Process

Table 4: Probabilities of touching sloping support resistance

t-statistic for $\mu_1 - \mu_2$, $t_{\alpha/2,\nu} = \{1.96, 2.576\}$for $\alpha = \{0.025, 0.005\}$								
p(resolution)	1	2	4	6	8	10	15	20
ups	5.558	5.378	3.786	3.212	3.753	1.955	0.237	-1.466
downs	5.255	4.553	3.052	3.523	1.556	1.875	-0.679	-3.973

Table 5: Test statistic for $\mu_1 - \mu_2$ sloping support / resistance

We see that for smaller resolutions the probability of touching the resistance levels (ups) are significantly higher in the observed time series than for the Wiener process. For the support levels (downs) it is also significant for the smaller resolutions (up till 6), but for smaller it is not significant, and even for the resolution 20, the probability is significantly higher for the Wiener process.

14.3 Support and resistance in moving average lines

Exponential moving average 200 is believed by many technical analyst to be a key signal for support and resistance level. I have in this test calculated the landmarks for different resolutions p, and counted the number of landmarks / turningpoints that are in the vincinity of the exponential moving average line. I have split the landmarks in *ups* and *downs,* (local maximas and local minimas) such that the *ups* landmarks touching the exponential moving average line indicates that the line acts as a resistance line, and for the *downs* the exponential moving average line serves as a support line. I have divided the number of reoccurences to the exponential moving average level by the total number of respectively local minimas and local maximas.

14.3.1 Eurodollar

Below are the probability of a turning point of resolution p is touching the exponential moving average line, for the eurodollar spot price. The exponential moving average, EMA(200) is given an errorband of 0.00025.

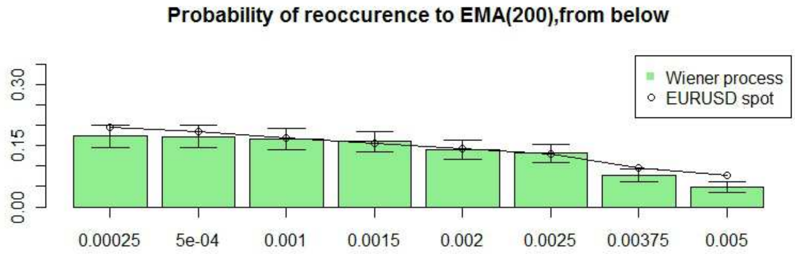
mean probability of reoccurence to EMA(200)								
p (resolution)	1	2	4	6	8	10	15	20
Observed *ups*	0.195	0.184	0.168	0.157	0.142	0.130	0.094	0.077
Wiener process *ups*	0.173	0.172	0.167	0.160	0.140	0.131	0.076	0.050

(a) Probability of reoccurence from below(*ups*)

mean probability of reoccurence to EMA(200)								
p (resolution)	1	2	4	6	8	10	15	20
Observed *downs*	0.170	0.161	0.144	0.144	0.131	0.117	0.096	0.079
Wiener process *downs*	0.209	0.207	0.207	0.179	0.173	0.151	0.109	0.049

(b) Probability of reoccurence from above(*downs)*

Table 6: Probability of reoccurence to the EMA(200) line



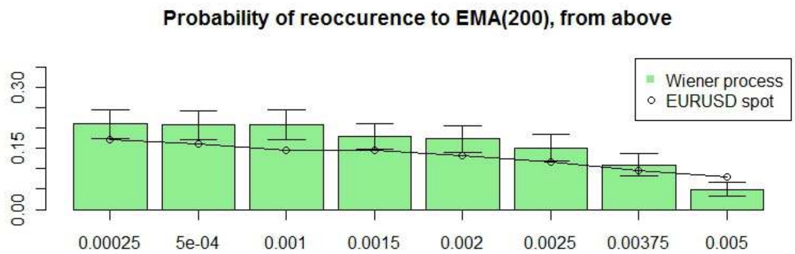


Figure 21: Reoccuring probability EMA(200) EURUSD

14.3.2 DAX futures

Below are the probabilities that the turning points of resolution p are touching the exponential moving average line, for the DAX futures price. The exponential moving average EMA(200), is given a error band of 1 point.

mean probability of reoccurence to EMA(200)								
p (resolution)	1	2	4	6	8	10	15	20
Observed *ups*	0.269	0.234	0.178	0.128	0.090	0.062	0.032	0.021
Wiener process *ups*	0.215	0.210	0.190	0.148	0.095	0.055	0.013	0.019

(a) DAX futures

mean probability of reoccurence to EMA(200)								
p (resolution)	1	2	4	6	8	10	15	20
Observed *downs*	0.281	0.245	0.181	0.120	0.079	0.057	0.040	0.046
Wiener process *downs*	0.216	0.212	0.188	0.148	0.101	0.055	0.014	0.015

(b) Wiener Process

Table 7: Test statistic support resistance EMA

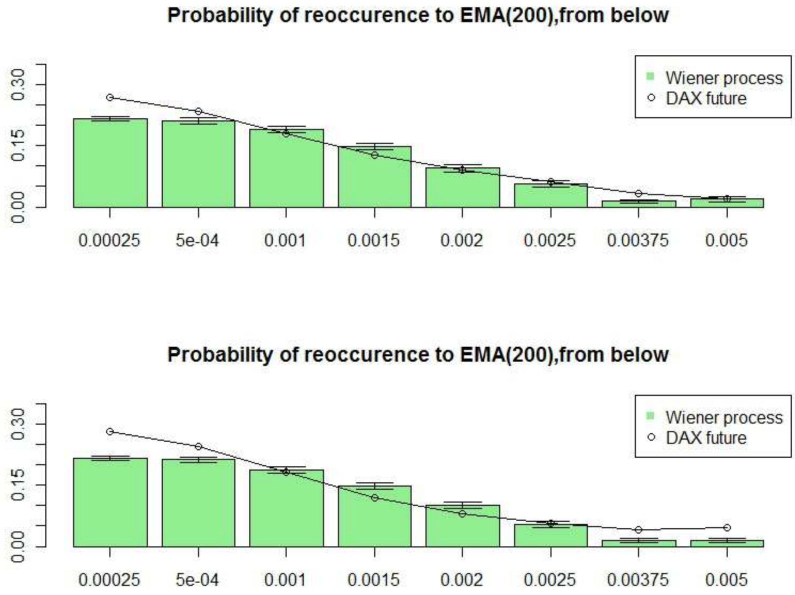


Figure 22: Reoccuring probability EMA(200) DAX

14.3.3 Inference about the exponential moving average EMA(200)

In comparing the observed prices to the Wiener process, we can see that from figure 19 that it is no evidence that the EMA(200) would be a significant indicator for turning points around the EMA(200) level for the EURUSD. For the DAX futures however it seems to be something for small and big resolution of the turning points. We'll test if there is something to it by the two-sample t-test, where μ_1 =observed data, and μ_2 = Wiener process

t-statistic for $\mu_1 - \mu_2$, $t_{\alpha/2,\nu} = \{1.96, 2.576\}$for $\alpha = \{0.025, 0.005\}$								
p(resolution)	1	2	4	6	8	10	15	20
ups	0.750	0.427	-0.0381	-0.112	0.070	-0.051	.985	1.625
downs	-1.334	-1.579	-2.159	-1.312	-1.557	-1.286	-0.554	1.726

(a) Test-statistic for EURUSD

t-statistic for $\mu_1 - \mu_2$, $t_{\alpha/2,\nu} = \{1.96, 2.576\}$for $\alpha = \{0.025, 0.005\}$								
p(resolution)	1	2	4	6	8	10	15	20
ups	11.638	4.626	-2.032	-3.163	-0.727	1.195	3.842	0.445
downs	13.945	6.720	-1.232	-4.454	-3.471	0.362	4.842	4.853

(b) Test-statistic for DAX future

Table 8: Test statistic for $\mu_1 - \mu_2$

15 Trading strategies

We have implemented intraday trading strategies that's based on the support and resistance levels. The support and resistance levels are predicted by calculating the support and resistance levels before 12:00:00 am, and then buying when "touching" the support levels, and selling short at the resistance levels. The strategy is also supported by stop loss and take profit arguments. This stop loss and take profit arguments are based on what is defined in literature as the most sensible values for each respectively, and is varying mostly in context to the difference between the support and resistance levels. By testing a strategy by backtesting it, we can find out if it deviates, and in case how much it deviates from the *efficient market hypothesis.* The deviation from the *efficient market hypothesis* is known as the edge of the strategy, in other words, what is the probability of making consistently more money than in a efficient market environment.

- Selling short is when selling an asset / security without owning it in the first place. There is sometimes introduced short selling restrictions on stocks, but on futures taking a long (buying asset) or a short (selling asset) position is treated the same.
- Stop loss is a parameter given in the strategy, and is the maximum loss one is willing to take on each trade. The stop loss could also be considered as a level at where the trading signal that initiated the trade was a "false signal".
- Take profit is another parameter in the strategy, where the "goal" for the trade was met, and the position is closed and profit is made.
- Backtesting is when running the strategy on historical datasets, in a same way one would run the strategy real-time. So it's not allowed to 'peak' at the future, so that each trading desicion at time t can only be based on $S_t\,, t \in [0, t]$.

15.1 Efficient Market Hypothesis

The efficient market hypothesis, was introduced by Eugene Fama in the early 70's and states that the market consists of only logical investors, and that all the market infomation available to all inverstors, and is already discounted in the price. There exist no opportunities to make consistently more than the risk free interest rate in the market. These opportunities would be what is called arbitrage opportunities. Mathematically this means that the expected portfolio wealth at time t is equal to the discounted wealth at time t, and that the stochastic process follows the martingale property.

$$V_0 = initial\, wealth \tag{46}$$

$$V_t = wealth\, at\, time\, t \tag{47}$$

$$\tilde{V}_t = discounted\, wealth\, at\, time\, t \tag{48}$$

then

$$E\left[\tilde{V}_t\right] = V_0 \tag{49}$$

an arbitrage opportunity would be if:

$$p\left(\tilde{V}_t > V_0\right) > 0 \tag{50}$$

and

$$p\left(\tilde{V}_t \geq 0\right) = 1 \ \ , V_0 = 0 \tag{51}$$

This indicates that there is a positive probability that the discounted wealth process V_t is greater than the initial wealth. And equation (52) indicates that wealth at time is greater or equal to zero a.s, given that initial wealth is zero.

- $h.r = p\,(trading\, signal\, is\, correct)$, hitting rate
- $prof$ =average profit when profitable
- $loss$ =average loss when loosing

$$E\left[(h.r\, prof) + ((1 - h.r)\ loss)\right] = 0 \tag{52}$$

Now we can estimate how much this deviates from the *efficient market hypothesis* by inputting the $prof$ and $loss$ from the backtest, and calculate the theoretical $h.r$ and compare it with the $h.r$ from the backtest.

15.2 Efficiency measures

A trading strategy or a portfolio is often measured in a standardized way with the *Sharpe ratio*[16]. The Sharpe ratio was introduced by William F. Sharpe, and is a measure of expected return over variance. The measure was introduced as a measure of mutual funds, and is now considered to be a branch standard for measuring performance of funds, and also performance of strategies.

[16] Reference 9

15.3 Ranging strategy

The ranging strategy is based on flat support and resistance levels, and as the name suggests, it is trading the range between the support and resistance. The support and resistance levels will be the maximum and minimum prices up to 12:00:00 am. Then if the price returns to the support level (minimum value up to 12 am.) we take a long position. And if the price returns to the resistance level, we take a short position. The position will now be closed at the opposite side of the range, so for a long position, the take profit level will be the resistance level, and for the short position, the take profit level will be the support level. The stop loss for the long position will be in our backtest the support level - 2 points / pips and for the short position the resistance level + 2 points/ pips. Points for the DAX futures, and pips for the eurodollar, pips is the minimum change in the eurodollar. If the price process goes above the minimum or the maximum, then there is done no trades this day, since the support and resistance levels are no longer valid.

Algorithm 6 Algorithm for ranging strategy

Inputs: max_open_12, min_open_12, trade=True, buy, sell, pnl, m = 1, stop_loss

1. for n in 1:T
2. while trade == True
3. if S_n =max_open_12, $sell_m = S_n$
4. if S_n >max_open_12 + stop_loss, $buy_m = S_n$, $pnl_m = sell_m - buy_m$, $m = m + 1$
5. if S_n =min_open_12, $buy_m = S_n$, $pnl_m = sell_m - buy_m$, $m = m + 1$
6. if S_n <min_open_12 - stop_loss, $sell_m = S_n$, $pnl_m = sell_m - buy_m$, $m = m + 1$
7. if S_n >max_open_12 + stop_loss or S_n <min_open_12, trade = False
8. end return pnl

From the performance plot below we see that the strategy is loosing in the over the dataset available. This could be due to the trading costs, or randomness in the signals.

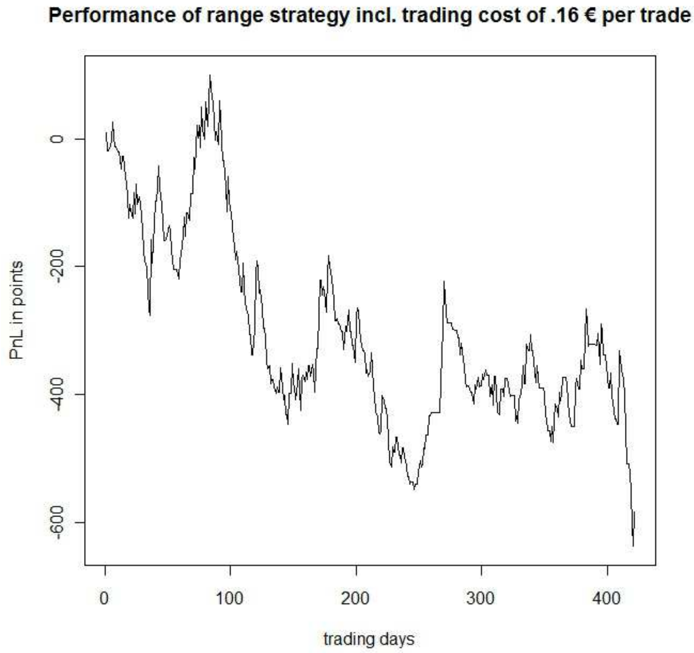


Figure 23: Performance of ranging strategy

15.4 Breakout strategy

The breakout strategy is in a similar fasion as the ranging strategy based on the support and resistance levels, but will take the opposite direction if the treshold levels are broken. eg. if the price process goes above the resistance level, one would take a long position (buy the security), and if the price process goes below the support level, one would take a short position (sell the security). A stop loss input in the breakout strategy would typically be the support or resistance level it has just broken through.

Algorithm 7 Algorithm for Breakout strategy

Inputs: max_open_12, min_open_12, trade=True, buy, sell, pnl, m = 1

1. for n in 1: length S_t
2. while trade == True
3. if S_n >max_open_12, $buy_m = S_n$
4. if S_n >max_open_12 + (max_open_12 - min_open_12) $sell_m = S_n$, $pnl_m = sell_m - buy_m$, $m = m + 1$, trade = False
5. if $S_n - buy_m < -2$, $sell_m = S_n$, $pnl_m = sell_m - buy_m$, $m = m + 1$
6. if S_n <min_open_12, $sell_m = S_n$
7. if S_n <min_open_12 - (max_open_12 - min_open_12) $buy_m = S_n$, $pnl_m = sell_m - buy_m$, $m = m + 1$, trade = False
8. if $S_n - sell_m > 2$, $buy_m = S_n$, $pnl_m = sell_m - buy_m$, $m = m + 1$
9. end return pnl

From the performance chart for the breakout strategy we see that this strategy is also loosing over the available data set period, but here also with quite big variations in the PNL (profit and loss).

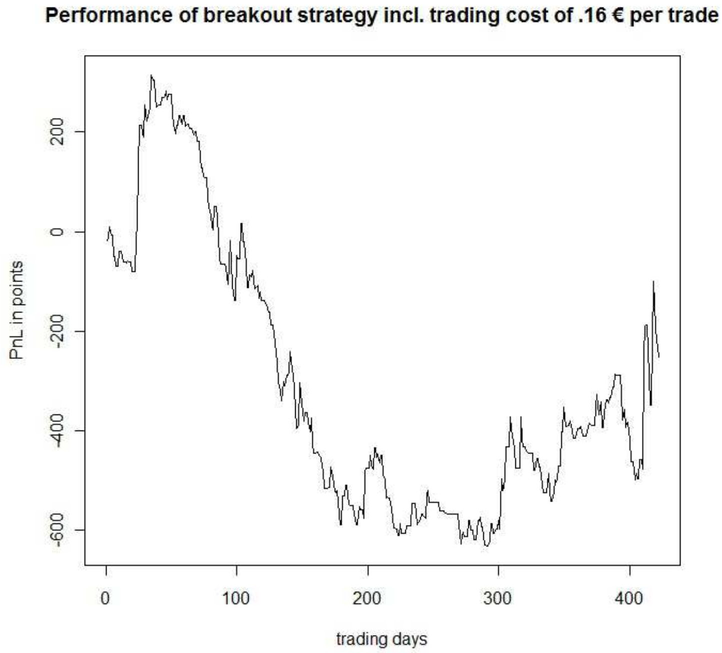


Figure 24: Performance of breakout strategy

15.5 Range and breakout strategy

The range and breakout strategy is both a ranging strategy and a breakout strategy based on the same support and resistance levels. The strategy will trade the range if the price is fluctuating within the support and resistance, and will take a directional position if the support or resistance levels are broken.

Algorithm 8 Algorithm for Breakout strategy

Inputs: max_open_12, min_open_12, trade=True, buy, sell, pnl, m = 1

1. for n in 1: length S_t
2. while trade == True
3. if S_n >max_open_12, $buy_m = S_n$
4. if S_n >max_open_12 + (max_open_12 - min_open_12) $sell_m = S_n$, $pnl_m = sell_m - buy_m$, $m = m + 1$, trade = False
5. if $S_n - buy_m < -2$, $sell_m = S_n$, $pnl_m = sell_m - buy_m$, $m = m + 1$
6. if S_n <min_open_12, $sell_m = S_n$
7. if S_n <min_open_12 - (max_open_12 - min_open_12) $buy_m = S_n$, $pnl_m = sell_m - buy_m$, $m = m + 1$, trade = False
8. if $S_n - sell_m > 2$, $buy_m = S_n$, $pnl_m = sell_m - buy_m$, $m = m + 1$
9. end return pnl

From the performance chart for the range and breakout strategy, we see that this strategy is gaining over the available data set period. It is however not much, and the variation is also quite big. It can also be that it is profitable by randomness, and actually not performing better than the previous strategies.

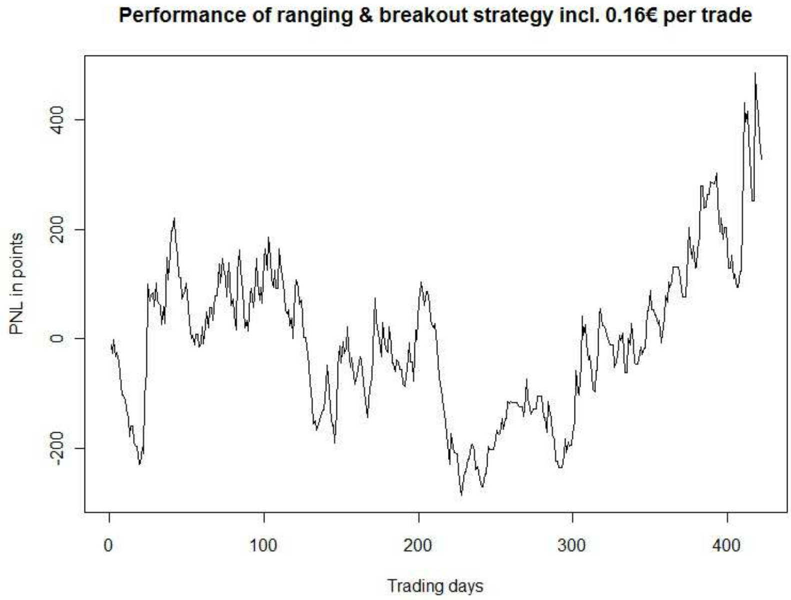


Figure 25: Performance of range and breakout strategy

15.6 Exponential moving average strategy

This strategy will use the Exponential moving average(EMA) lines as a support and resistance line, and perform one trade per day based on the signals given. If the price reaches the EMA- error band, from below, the strategy will take a short position. If the price reaches the EMA + error band from above, the strategy will take a long position.

Algorithm 9 Algorithm for Exponential moving average strategy

Inputs: EMA(200), S_t, $error.margin = 1$, trade=True, buy, sell, pnl, m = 1

1. for n in 201: length S_t
2. while trade == True
3. if $S_n > EMA(200)_n - error.margin$ & $S_{n-1} < EMA(200)_{n-1}$ $sell_m = S_n$
4. if $S_n > EMA(200)_m + error.margin$,$sell_m = S_n$, $pnl_m = sell_m - buy_m$, $m = m + 1$, trade = False
5. if $S_n - sell_m > 2$, $buy_m = S_n$, $pnl_m = sell_m - buy_m$, $m = m + 1$
6. if $S_n < EMA(200)_n + error.margin$ & $S_{n-1} > EMA(200)_{n-1}$ $buy_m = S_n$
7. if $S_n < EMA(200)_m - error.margin$,$buy_m = S_n$, $pnl_m = sell_m - buy_m$, $m = m + 1$, trade = False
8. if $S_n - buy_m < -2$, $sell_m = S_n$, $pnl_m = sell_m - buy_m$, $m = m + 1$
9. end return pnl

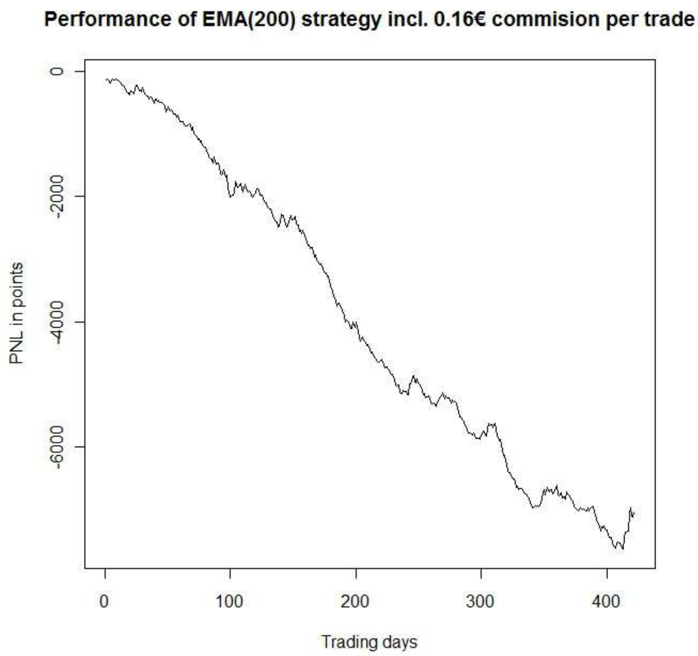


Figure 26: Performance of EMA(200) as support & resistance strategy

The performance of this strategy is really poor. This comes from that the strategy is making a lot of trades, due to many trading signals generated from the EMA(200) level. For each transaction we are deducing the commision of 0.16€. It could also be that the error margin is to tight, so that the strategy is exiting the position because of "noise". When we added a constraint that the strategy would only close a profitable position, or close at end of day, the strategy went from loosing about 7000 points, to making about 600 points, but with a Sharpe ratio below 0.5 indicating that the strategy is very volatile.

16 Extreme landmarks

The concept of the extreme landmarks is that, with in the landmarks that are calculated, there are certain with more extreme movements over a very short period of time. We want to test if there exist some jumps bigger than the given resolution p from the landmark model times a given λ, followed by a jump in the opposite direction of the same size or bigger. Mathematically we will be looking at landmarks, where,

$$LM_{t-1} < LM_t > LM_{t+1} \tag{53}$$

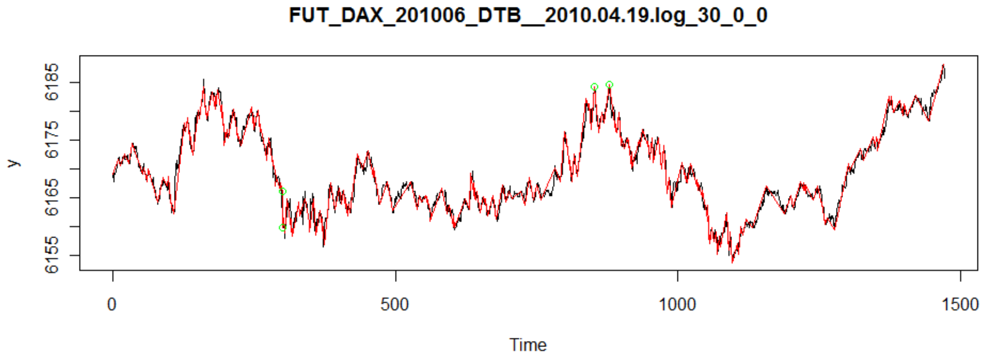
or

$$LM_{t-1} > LM_t < LM_{t+1} \tag{54}$$

and

$$\begin{aligned} |LM_{t-1} - LM_t| &> p\lambda \quad (55)\\ |LM_t - LM_{t+1}| &> p\lambda \quad (56) \end{aligned}$$

for some fixed λ and p as resolution from the original landmark calculations.



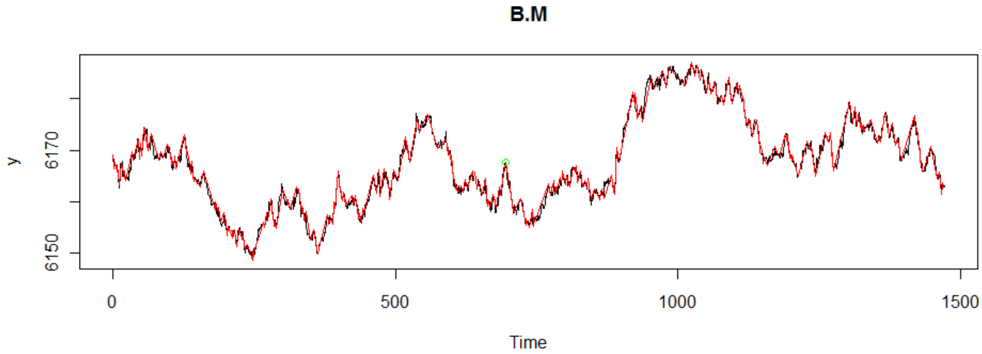


Figure 27: Extreme landmarks
The landmarks detected as being extreme has movement bigger or equal to $p\lambda$ in one direction, followed by a movement of same size in the opposite direction.

Observed (mean)				Wiener process (mean)			
P / λ	2	3	4	P / λ	2	3	4
2	12.11	2.03	0.48	2	10.79	0.95	0.12
4	1.89	0.18	0	4	2.59	0.22	
6	0.68	0.02	0	6	1.17	0.11	0

Observed (st.dev)				Wiener process (st.dev)			
P / λ	2	3	4	P / λ	2	3	4
2	0.38	0.14	0.06	2	0.28	0.07	0.03
4	0.12	0.03	0	4	0.13	0.03	0
6	0.06	0.01	0	6	0.08	0.02	0

Table 9: Extreme landmarks

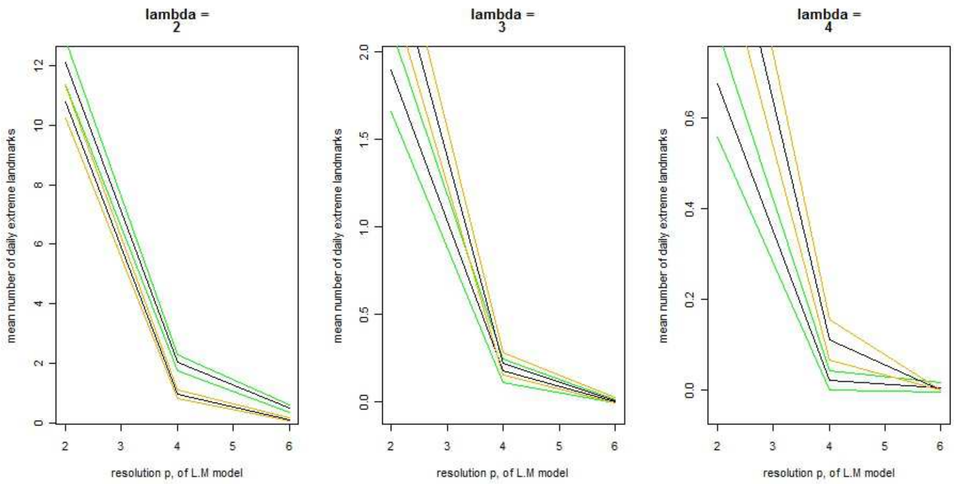


Figure 28: Extreme landmarks
The mean number of extreme landmarks

t-statistic for $\mu_1 - \mu_2$, $t_{\alpha/2,\nu} = \{1.96, 2.576\}$for $\alpha = \{0.025, 0.005\}$			
p(resolution) \ λ	2	3	4
2	-8.544	-1.829	-3.706
4	-0.715	0.962	1.885
6	3.913	0.182	1.138

Table 10: t-tests of extreme landmarks
We see from the t-tests that for $p = 2$, and $\lambda = 2$ the number of landmarks that are twice the size of the resolution and bounces back with the same size are occuring significantly more frequent in the Wiener process than in the observed data.

From the t-tests of the extreme landmarks, we see that for $p = 2$, the frequency of landmarks that follows (53-56) have significantly higher occurence for $\lambda = 2$ and 4 in the Wiener processes. For resolution $p = 6$ we have significantly higher frequency of landmarks in the observed dataset for $\lambda = 2$.

Part VI
Concluding remarks

We have developed a new way of estimating intraday volatility, by using local minimas and maximas, in a time series. The result is an unbiased estimator based on the assumption that the market follows a Wiener process without drift. In the efficiency section we see that the efficiency is one fourth of the classical estimator. This is not very ideal, since it indicate that we would need four times the length of the datasample to be able to get the same certainty of the variance. There are however more ways to further develop from this volatility estimator, that could still make it very interesting. By in example changing the assumption of that the market behavies in a Geometric Brownian motion rather than in a Wiener process without drift. In section 9.5 till 9.9 we briefly introduce an extension of the estimator to include a jump size and intensity estimate. The extension is only briefly introduced, and will be left for later work.

In Part V we have done several tests to check if there are certain levels within the day that could work as either support or resistance levels. We found in the entropy section significant evidence of certain levels that have a much higher hitting probability in the observed datasets than in the Wiener process. Tests were also done for sloping levels, and there are significantly higher reoccurence probabilities to the levels we have found using algorithm 5. In section 15 we have defined trading strategies based on the founds in the previous sections, but don't find any really profitable trading strategies based on the signals. This could be because of the weakness of the signal, or the prediction power of the signal since we are not looking forward in the backtesting of the trading strategies. We have also deduced an approximate of commision on each trade, and this could also be a significant reason for the results in the trading strategies section.

In section 16. we have looked at landmarks that are bigger than a factor apart from the other landmarks, and then suddenly bounces back with the same or greater size. We have found little or no evidence in the tests that there exists jumps that behaves in this characteristics.

Part VII

References

1. Chang-Shing Perng, Haixun Wang, Sylvia R. Zhang, D. Stott Parker, Landmarks: A New Model for Similarity-Based Pattern Querying in Time Series Databases. Proc. 16th Int. Conf. on Data Engineering, ICDE, pp. 33-42(2000).

2. Martin Pring, Technical Analysis Explained : The Successful Investor's Guide to Spotting Investment Trends and Turning Points, Fourth ed., 2002, McGraw-Hill

3. Claude E. Shannon, A Mathematical Theory of Communication, The Bell System Technical Journal, Vol. 27, pp. 379–423, 623–656, July, October, 1948.

4. Jay L. Devore, Kenneth N. Berk, Modern Mathematical Statistics with Applications, 2007, Duxbury

5. Bernt Øksendal, Stochastic Differential Equations, 2003. Springer. (6th Edition)

6. Andrew W. Lo, Harry Mamaysky, Jiang Wang, Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation, The Journal of Finance, Vol. LV, No. 4 August 2000.

7. Mark B. Garman and Michael J. Klass, *University of California, Berkeley,* On The Estimation of Security Price Volatility From Historical Data

8. Rama Cont and Peter Tankov, Financial Modelling With Jump Processes, 2004, Chapman & Hall/CRC

9. William F. Sharpe, The Sharpe Ratio, Reprinted from The Journal of Portfolio Management, Fall 1994